AD-A094 625 CANYON RESEARCH GROUP INC WESTLAKE VILLAGE CALIF F/G 5/9
SIMULATOR TRAINING REQUIREMENTS AND EFFECTIVENESS STUDY (STRES)--ETC(U)
JAN 81 W W PROPHET, J B SHELNUTT, W D SPEARS F33615-77-C-0067

UNCLASSIFIED AFHRL-TR-80-37 NL

1 OF 2

AD-
A094625

AFHRL-TR-80-37

AIR FORCE

HUMAN RESOURCES

LABORATORY

AD A094625

SIMULATOR TRAINING REQUIREMENTS AND EFFECTIVENESS STUDY (STRES): FUTURE RESEARCH PLANS

By

Wallace W. Prophet
Jack B. Shelnutt
William D. Spears
Seville Research Corporation
400 Plaza Building
Pensacola, Florida 32505

LOGISTICS AND TECHNICAL TRAINING DIVISION
Logistics Research Branch
Wright-Patterson Air Force Base, Ohio 45433

January 1981

Final Report

Approved for public release; distribution unlimited.

DDC FILE COPY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

81 2 04 071

NOTICE

When U.S. Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This final report was submitted by Seville Research Corporation, 400 Plaza Building, Pensacola, Florida 32505, under Contract F33615-77-C-0067, Project 1710, with the Logistics and Technical Training Division, Air Force Human Resources Laboratory (AFSC), Wright-Patterson Air Force Base, Ohio 45433. Mr. Bertram W. Cream was the Contract Monitor for the Laboratory.

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

ROSS L. MORGAN, Technical Director
Logistics and Technical Training Division

RONALD W. TERRY, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFHRL-TR-80-37	2. GOVT ACCESSION NO.: AD-A094 625	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): SIMULATOR TRAINING REQUIREMENTS AND EFFECTIVENESS STUDY (STRES): FUTURE RESEARCH PLANS		5. TYPE OF REPORT & PERIOD COVERED: Final
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Wallace W. Prophet, Jack B. Shelnutt, William D. Spears		8. CONTRACT OR GRANT NUMBER(s): F33615-77-C-0067
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Seville Research Corporation, 400 Plaza Building, Pensacola, Florida 32505		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62205F, 17100342
11. CONTROLLING OFFICE NAME AND ADDRESS: HQ Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235		12. REPORT DATE: January 1981
		13. NUMBER OF PAGES: 178
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Logistics and Technical Training Division, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base, Ohio 45433		15. SECURITY CLASS. (of this report): Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

simulation research	aircrew training devices	training technology	aircrew performance
simulator research	simulator technology	training effectiveness	
training research	simulator utilization	transfer of training	
research program	aircrew training	learning	
training device research	flight training	fidelity	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report, prepared for Air Force personnel responsible for the development and execution of future programs of research on the design, development, and use of aircrew training devices (ATDs), presents topics to be considered in the development of future ATD research programs. The topics were developed in the course of the Simulator Training Requirements and Effectiveness Study (STRES). One objective of STRES was to identify gaps in the technology base that underlies the development and use of ATDs. Research needs were identified on the basis of reviews of the simulation research literature; observations and interviews conducted at a wide variety of simulator training agencies; interviews and discussions with simulator research, procurement, management, and design personnel; and comparisons of suggested topics with plans of various ATD research agencies.

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Item 20 Continued:

The report presents a basic listing of 111 potential research topics. Detailed research plans were developed for 21 of these topics of highest priority. Each plan includes: (a) problem statement; (b) research overview; (c) analytic requirements; (d) suggested experimental methodology; (e) subjects requirements; (f) data collection and analysis plans; (g) facilities requirements; and (h) projected schedule and manpower requirements. Less detailed plans are presented for nine additional topics.

PREFACE

This report describes a portion of a study of training through simulation in the U.S. Air Force. It is one of seven technical reports prepared for the Air Force Human Resources Laboratory, Logistics and Technical Training Division, under Contract F33615-77-C-0067, Simulator Training Requirements and Effectiveness Study (STRES). The remaining six reports are identified in Chapter II of this document. The reports cover work performed from August 1977 through January 1980.

The work was performed by a team made up of Canyon Research Group, Inc; Seville Research Corporation; and United Airlines Flight Training Center. Canyon Research Group, Inc. was the prime contractor; Mr. Clarence A. Semple served as the Program Manager. The Seville Research Corporation effort was headed by Dr. Paul W. Caro. The United Airlines effort was headed initially by Mr. Dale L. Seay and subsequently by Mr. Kenneth E. Allbee.

Mr. Bertram W. Cream was the AFHRL/LR Program Manager. Other key members of the AFHRL/LR technical team included Dr. Gary Klein and Dr. Thomas Eggemeier. A tri-service STRES Advisory Team participated in guiding and monitoring the work performed during this contract to assure its operational relevance and utility. Organizations participating in the Advisory Team were:

Headquarters, USAF
Headquarters, Air Training Command
Headquarters, Military Airlift Command
Headquarters, Aerospace Defense Command
Headquarters, Tactical Air Command
Headquarters, Air Force Systems Command
Headquarters, Strategic Air Command
Tactical Air Warfare Center
Air Force Manpower and Personnel Center
Air Force Test and Evaluation Center
USAF Aeronautical Systems Division
Air Force Human Resources Laboratory
Air Force Office of Scientific Research
Army Research Institute for the Behavioral and Social Sciences
United States Navy Training Analysis and Evaluation Group

The authors wish to express their gratitude to the hundred of people in the United States Air Force, Navy, Army, Coast Guard, NASA, FAA and industry who contributed to this program by participating in interviews and technical discussions during program data collection.

TABLE OF CONTENTS

PRECEDING PAGE BLANK-NOT FILMED

LIST OF TABLES

I. INTRODUCTION

BACKGROUND

The U.S. military services have been users of flight training devices and simulators for over half a century. These training media, known collectively as aircrew training devices (ATDs), include cockpit familiarization and procedures trainers, part-task trainers, mission trainers, operational flight trainers, and weapons systems trainers. In recent years, use of ATDs has increased to the point that the devices represent major aircrew training resources, and the effectiveness and efficiency of their design and use are matters of continuing concern.

The STRES Project

In response to this concern, the U.S. Air Force, in cooperation with the other military services, undertook a programmatic study of factors involved in ATD design and use. This program was titled Simulator Training Requirements and Effectiveness Study (STRES). The general objectives of STRES are to define, describe, collect, analyze, and document information bearing on the cost and training effectiveness of flight simulators. Topic areas covered in the program include fidelity, instructional support features, simulator utilization, life cycle costs, and worth of ownership. Products of the program are intended for use by those who manage and use simulators for training, those who evaluate simulator requirements, and those who design, procure, and maintain these devices. In addition, STRES results are intended to assist those who conduct and manage research on ATDs and ATD training.

The overall plan of the STRES program calls for a four-phase effort. Phase I, which has been completed by the Air Force, structured the general problem so that operationally relevant ATD training issues could be addressed on a priority basis. Phase II, which is the subject of this and related reports, was a 29-month effort that involved collecting, integrating, and presenting available information on the design, use, life cycle cost, and worth of ownership of ATDs. Phase III, to be conducted in the future, will involve the research activities necessary to answer important questions about ATDs and training effectiveness that cannot be answered with assurance using information presently available. The research to be conducted in Phase III will be the responsibility of AFHRL/OT, and the research plans presented in the present report are topics suggested as input to Phase III. Building on the work done in Phases II and III, Phase IV will be an Air Force effort to integrate findings and publish new information as it becomes available.

Phase II Objectives

Five general technical objectives emerged for Phase II of the STRES program. These were:

1. Develop criteria for matching training requirements with ATD features and for degree of fidelity required to meet requirements efficiently and effectively; and

2. Define principles of effective and efficient use of ATDs to accomplish specified training requirements; and

3. Develop criteria for matching instructional support features with specified training requirements; and

4. Develop models of factors influencing the worth of ownership and life cycle costs of ATDs; and

5. Develop a plan for future Air Force training research and development regarding ATD features, ATD use, and ATD training effectiveness where present knowledge or technology is inadequate.

PURPOSE OF THIS REPORT

The purpose of this report is to present the findings of the Phase II effort with respect to the fifth and final Phase II objective--development of research plans for future Air Force training R&D regarding ATD features, ATD use, and ATD training effectiveness where present knowledge or technology is inadequate.

Context of this Report

This report on future research is one of seven documents published to report results of the STRES project. The reports are as follows:

Caro, P.W., Shelnutt, J.B., & Spears, W.D. Aircrew Training Devices: Utilization. AFHRL-TR-80-35. Wright-Patterson AFB, OH: Logistics and Technical Training Division, Air Force Human Resources Laboratory, January 1981.

Semple, C.A., Hennessy, R.T., Sanders, M.S., Cross, B.K., Beith, B.H., & McCauley, M.E. Aircrew Training Devices: Fidelity Features. AFHRL-TR-80-36. Wright-Patterson AFB, OH: Logistics and Technical Training Division, Air Force Human Resources Laboratory, January 1981.

Semple, C.A., Cotton, J.C., & Sullivan, D.J. Aircrew Training Devices: Instructional Support Features. AFHRL-TR-80-58. Wright-Patterson AFB, OH: Logistics and Technical Training Division, Air Force Human Resources Laboratory, January 1981.

Allbee, K.E., & Semple, C.A. Aircrew Training Devices: Life Cycle Cost and Worth of Ownership. AFHRL-TR-80-34. Wright-Patterson AFB, OH: Logistics and Technical Training Division, Air Force Human Resources Laboratory, January 1981.

Prophet, W.W., Shelnutt, J.B., & Spears, W.D. Simulator Training Requirements and Effectiveness Study (STRES): Future Research Plans. AFHRL-TR-80-37. Wright-Patterson AFB, OH: Logistics and Technical Training Division, Air Force Human Resources Laboratory, January 1981.

Spears, W.D., Sheppard, H.J., Roush, M.D., & Richetti, C.L. Simulator Training Requirements and Effectiveness Study (STRES): Abstract Bibliography. AFHRL-TR-80-38. Wright-Patterson AFB, OH: Logistics and Technical Training Division, Air Force Human Resources Laboratory, January 1981.

Semple, C.A. Simulator Training Requirements and Effectiveness Study (STRES): Executive Summary. AFHRL-TR-80-63. Wright-Patterson AFB, OH: Logistics and Technical Training Division, Air Force Human Resources Laboratory, January 1981.

The first four reports address the first four objectives of the Phase II effort. They provide guidance concerning the design and utilization of ATDs based on the analysis of currently available technological information. The present report was conceived in coordination with the first four to address gaps in current ATD utilization and design technology that restrict the maximization of the effectiveness and efficiency of ATD training in the Air Force. Thus, this report supplements the first four reports by providing research plans relevant to the principal issues for which proper guidance was not found. For details concerning the literature reviewed and the various observations made and information gathered during STRES, the reader is referred to the other STRES reports. While certain aspects of the information covered in those reports are summarized and discussed in the present report, no attempt is made to cover their contents completely. For ease of reference in the present report, the other STRES reports will be referred to by abbreviated title (Fidelity, Instructional Features, Utilization, or Cost). The other STRES reports contain a variety of references to research needs, both major and minor. For obvious practical reasons, the research plans presented in the present report do not provide exhaustive treatment of all gaps in ATD utilization and design technology. However, they do address the major problems identified during Phase II of STRES.

Focus of the Research Plans

The research plans focus on topics that are responsive to Air Force needs in the simulator design and utilization areas, but only on needs that are amenable to solution through the conduct of experimental training research. Thus, the plans do not address certain critical issues that are inappropriate for experimental training research, for example, certain management issues and problems, nor do they address issues that are primarily engineering in nature. However, many such issues are discussed in the other STRES reports. More specifically, the focus of the plans was restricted to research that addressed information needs for relating aircrew training effectiveness and efficiency to (1) ATD fidelity factors, (2) ATD instructional support features, and (3) ATD utilization factors. In all such research areas identified, cost factors represent an underlying concern, even though cost itself is not a direct experimental topic. This focus was further sharpened in terms of a set of "high value" operational tasks and issues specified in the contract Statement of Work (SOW) for which ATD training was believed to offer potential benefits, and for which information was lacking (or at least not integrated) regarding ATD design, use, and worth. The SOW high value tasks are:

- Individual and formation takeoff and landing;
- Close formation flight and trail formation, both close and extended;
- Aerobatics;
- Spin, stall, and unusual attitude recognition, prevention, and recovery;
- Low-level, terrain-following flight;
- Air refueling;
- Air-to-air combat (both guns and missiles); and
- Air-to-ground weapons delivery.

REPORT ORGANIZATION

This report consists of three chapters and four appendices. The first chapter is this introduction. In addition to the background section, Chapter I also contains a discussion of certain basic considerations that underlay the selection of research topics for the detailed research plans.

Chapter II describes the approach that was employed to develop the research plans. Chapter III presents an integrative overview of the research topics that were selected as subjects for detailed research plans.

The appendices list the sites visited by the STRES team, as well as present the research topics that were suggested during Phase II. Detailed descriptions of the content of each appendix and the approach used to generate its content are presented in Chapter II.

BASIC CONSIDERATIONS

The details of the research plans developed have been influenced by a number of considerations. Obviously, the basic purposes described in the contract work statement have been primary, but varying emphases have been provided by the evolutionary development of the STRES effort itself. While such interactive influence was expected and presumed, the nature of such influence is worth delineation as background for relating the research plan to (1) the general STRES findings and (2) the ongoing simulation research programs of various agencies.

Research and Future Impact

The development of the research plan necessarily involved the establishment of priorities among potential research topics. Such priorities are required because the resources that will be available to support research will be limited, and the resources available to support the development of the research plan itself were also limited. Therefore, a principal concern has been the identification of topics that will offer the greatest potential for the enhancement of future Air Force simulation programs.

The research topics that have been considered for possible inclusion in the research plan can be categorized, generally, as either relating to simulator utilization or to simulator technology. Of these two categories of future research, that related to simulator utilization, i.e., to aspects of simulator training programs or to the learning and transfer-of-training processes, is judged by the STRES team to offer more overall potential for significant program enhancement over the next 5-10 years. This is not to say, of course, that research related to the simulator hardware/software technology areas is not needed; it is emphatically needed, and the research topics described later in this document reflect this. However, the findings reported in the other STRES documents support the conclusion that, in terms of significant impact on future Air Force simulation programs, the greatest potential gain in benefits overall would seem to lie in improved utilization practices. Hence, this leads to emphasis on the utilization research area. The rationale for this assertion is described in following paragraphs.

It is obvious that devices of great potential capability can be misused so as to produce little or no training benefit, and that relatively unsophisticated devices can produce considerable training benefit if utilized in well designed device training programs. These facts have received increasing attention from simulation training researchers in recent years. At one level, this concern has dealt with a series of issues that are essentially of a management nature. For example, the extent to which managers require clear specification of training objectives, systematic procedures for the design of simulator training programs, adherence to program requirements, training for instructor personnel, proper maintenance of hardware/software, appropriate integration of simulators with other media, and similar "common sense" training management procedures, all relate to how effectively the simulator is utilized.[1] Underlying such management procedures, however, are a number of significant <u>training technology</u> questions that must be answered if utilization practices are to be enhanced. It is the seeking of answers to these questions that defines the nature of required utilization research. Such questions relate to development of better understanding of human learning in the simulator context, the cues required for effective simulator training, the nature of mediational processes, the use of feedback, and similar concerns that wittingly or unwittingly underlie simulator utilization.

<u>Overview of Status of Simulation in USAF</u>

The STRES effort has provided an extensive opportunity to examine Air Force simulator programs in terms of the equipment and systems available currently or to be available in the next decade, the personnel involved with simulation, and current operational simulator training programs. It has also allowed the opportunity to examine those factors against the backdrop of future Air Force training needs and against simulator research accomplished and planned for the future by both Air Force and non-Air Force agencies. The reader is referred to the other STRES reports for more extensive treatment of these subjects, but, as background for the research program described in the present report, the following summarization of STRES findings is offered with respect to utilization, simulation technology, currently planned research programs, and future simulator training needs.

---

[1]The procedural guidance outlined in the instructional systems development (ISD) methodology embraced by all the services provides a sound structure for identifying functional requirements and decision points that are basic to sound training development and management. However, ISD requires an underlying knowledge of human learning, a requirement that is becoming increasingly evident to training developers and managers alike. Thus, the research needed is not so much to better the ISD methodology as to develop the supporting technology and information base on which ISD rests.

Utilization. Examination of Air Force operational simulator utilization practices reveals a number of areas of need for research on improved utilization practices. As is noted in the STRES Utilization report, personnel responsible for Air Force ATD training activities generally did not appear to understand the complexity or nature of the training transfer process, tending to view it as a simple, almost mechanical outcome that is primarily dependent on stimulus and response fidelity. Most operational simulator training programs observed reflected the "aircraft" or "inflight training" model of instruction and, consequently, reflected little application of the learning-related principles on which simulation is (or should be) based. It is clear that this situation leaves significant room for the enhancement of future training effectiveness through better application of learning process information. This situation reflects, in part, inadequacies of dissemination of that which is known with reference to effective simulator utilization. While a good deal is presently known about effective simulator utilization, the situation also reflects certain inadequacies in our understanding of transfer, mediation, feedback, and similar learning-related factors that underlie simulation. Stated differently, there are still significant gaps in the technology and data bases that support simulator training.

As a result, the research program that follows reflects very strongly this concern with learning and other factors related to utilization. It may be observed that the emphases already emerging in the simulation research programs of the Air Force and the other services reflect very strongly a concern with increasing the effectiveness of operational applications of simulation in military flight training programs, particularly at the unit level.

Simulation technology. The Air Force has in hand or in procurement ATDs that represent highly advanced simulator technology. This technology is adequate to provide needed training capability in many areas. But, the utilization technology base must be expanded if the training potential of existing ATD technology is to be realized. However, there are other significant areas of ATD technology need in addition to utilization. Clearly, the increasing application of simulation to the various areas of visual task learning carries with it the need for significant visual technology developments. Many of these needs are being met by projects that are presently underway or are provided for in existing Air Force research plans. However, additional explorations of visual cue requirements for specific tasks are recommended here. Similarly, it is the position of the STRES team that certain research on motion cues and motion-force cueing mechanisms is required.

While there exist many instructional support features (ISF) on current simulators, the most needed research related to ISF would seem to be more of the utilization variety, e.g., how to use ISF more effectively with reference to instructional process control through feedback

and guidance, than of the simulation hardware/software type. However, there are a number of other training-related aspects of ATD design that warrant attention. For example, the general design of the instructor-operator station (IOS), the nature of information displays (e.g., visual system displays) for the instructor, and system requirements related to performance measurement systems are ATD technology areas in which future research should be beneficial.

Current research programs. As has been noted, there already exists in Air Force and other service research plans an increasing emphasis on research dealing with operational applications of simulation. Most past simulator research has focused on initial skill acquisition training in undergraduate pilot training (UPT) and aircraft transition training at the combat crew training (CCT) level, but the emphasis is moving toward research on operational combat skills and the needs of units in the continuation training (CT) area. One consequence of this change in emphasis has been the highlighting of complex visual simulation needs such as in the air-to-air combat area.

This emphasis on operational training is evident in the current simulation research plans for all the services, and it reflects their general concern with making ATD operational utilization more beneficial to units. It appears to the STRES team that utilization benefits will be further enhanced by an additional emphasis on learning process research, as previously described. Thus, the current service research program emphases in simulation technology (e.g., visual, motion, and performance measurement) and on operational training seem highly appropriate to the USAF simulation research needs as sensed by the STRES team, but such efforts should be supported by additional research on the learning process as it relates to simulation and simulator training.

Future simulation programs. The Air Force has initiated a series of major simulator procurement programs which represent a substantial investment. These programs will be aimed, largely, at the training needs of operational units. Thus, in examining the nature of future Air Force simulation programs, it is apparent that the major "frontier" will be in the use of simulation to support combat skills training at the CT level. The principal value of future simulation will not be so much to replace aircraft time, though this will continue to be of concern, but as a means of training critical combat skills that can only or best be trained through simulation approaches.

As a result of this CT emphasis, the major areas of simulation technology need will be those related to tactical skills, e.g., visual and motion system technology, performance measurement and combat skills assessment technology. However, as noted, full realization of training benefits from advances in these simulation technology areas will require advances in the training technology that is basic to both simulator design and use.

Technology Transfer

The research proposed here involves a program of experimental studies, many of which are programmatic in nature. Most programmatic research efforts involve a variety of specific projects that are structured to allow progress toward an overall goal of development of an information data base sufficient to allow effective application across a variety of situations. In the case of simulation research, or research on training generally, the actual applications to Air Force operational training problems typically will be made by personnel who, by and large, are not professional training technologists, education researchers, or education specialists; application will be largely in the hands of pilots.

In order for such "lay" applications of the various simulation and training technologies to be effective, the following conditions must be met. First, the technology base must be sound and derive from well conceived and well conducted research. Second, the technology must be stated in a form capable of successful implementation by nontechnological personnel, i.e., the technology must be transferable for application. One aspect of such transfer involves the development of a data or information base sufficient to support applications. However, even if these conditions are met, there exists a final condition or need; there must be some mechanism to effect the actual transfer.

One mechanism that can be most useful in the transfer process is the use of carefully developed demonstration programs. In fact, for this particular situation, it may reasonably be contended that demonstration is essential to effective technology transfer.

One of the principal and most pervasive problems noted in operational simulator training programs observed by the STRES team was the persistent use of the "aircraft training model" with its emphasis on physical and stimulus fidelity (as opposed to cue fidelity and mediation). Effective utilization of simulation often involves procedures that are at variance with the "aircraft training model," and, for this reason, the actual use made of an ATD under the aircraft model may be relatively ineffective. The demonstration program can make these differences in utilization models understood by applications personnel and can provide convincing evidence concerning the advantages of the "simulator training model." Such evidence is critical to acceptance and proper implementation of training practices and procedures that may be outside the flight experience or expectancies of the typical pilot.

Recent AFHRL studies of the efficacy of simulation[1] in the teaching of air-to-ground weapons delivery skills for A-10 pilots illustrate the

---

[1] It is of interest to note that this simulation exhibited a number of areas of simulator-aircraft variance in terms of physical and stimulus/response fidelity.

advantages of a demonstration, both for increasing user acceptance of simulation and for making manifest the details of the training technology involved. This particular demonstration was carried out in the context of an experimental research study, a fact which illustrates that a demonstration need not be outside the bounds of a research program. In any event, the potential of demonstration as a part of the technology transfer process is considerable. It is felt that with reference to more effective implementation of simulator training in the Air Force, it is a critical factor, one worthy of attention.

II. APPROACH

The approach employed to develop the research plans consisted of five major activities. These were: (1) the identification of data base needs and generation of potential research topics that addressed those needs; (2) comparison of the potential research topics with existing research plans of simulator research organizations to eliminate topics that duplicated or overlapped with the existing plans; (3) assessment of the relative priority of the remaining potential research topics; (4) review by the Air Force of the topics and their assessed priorities; and (5) selection of high priority topics for detailed consideration and preparation of research plans for those topics.

GENERATION OF THE LISTING OF POTENTIAL RESEARCH TOPICS

The objectives of the first activity were to determine major gaps that existed in the state of knowledge with respect to ATDs, i.e., the identification of data base needs, and to develop a listing of potential research topics that addressed those needs. As discussed in Chapter I, the identification process was focused in terms of the needs to which the STRES project was responsive and toward which future Air Force research could be addressed.

Sources of Information Concerning Data Base Needs

Information concerning data base needs was derived from three major sources. These were: (1) the technical literature; (2) site visit observations and interviews with personnel involved in aircrew training, ATD design and production, ATD procurement, and ATD training research and development; and (3) the analyses conducted as part of Phase II.

Literature review. To provide information for the entire Phase II effort, extensive computer and manual searches were made of the technical literature. More than 400 documents were identified during these searches as being useful for STRES purposes. These documents were reviewed to determine the current state of knowledge concerning ATD design and utilization. At this stage of the study, gaps were identified through analysis of training problems and ATD design issues reported in the documents. Furthermore, research was identified that was incomplete or inadequate for current data base needs.

In addition to this general review, evaluative abstracts were prepared for 200 of the articles as part of the overall Phase II effort. (These abstracts are published in a separate report). The abstracting process was particularly useful in the identification of data base needs, because it provided an opportunity for the in-depth analysis of research issues and of the adequacy with which they had been studied.

Site visits. A major aspect of the overall Phase II data collection effort involved site visits to a number of aircrew training, ATD manufacturing, ATD procurement, and aviation training R&D organizations. Organizations visited are shown in Appendix A.

The information sought during interviews varied with the type of site visited and with the responsibilities of individuals comprising each interview team. Personnel interviewed included those actively involved in ATD training and its management; persons involved in ATD procurement, testing, and support; command personnel; industry representatives; and persons actively involved with ATD research and development. The structure for these contacts was provided by interview guides. Copies of the guides used, as well as discussions of the information gathered, appear in the first four of the STRES reports.

As part of each interview, respondents were asked to identify ATD training or ATD design problems that they believed should be addressed. Personnel at simulator research organizations were also queried concerning their current research programs and plans for future research, and the facilities they had available or would have available in the near future for ATD research.

STRES project analyses. Analyses that were conducted to support the preparation of the Fidelity, Instructional Features, Utilization, and Cost reports also served in the identification of data deficiencies These analyses focused on the assessment of the information collected through the literature review and site visits and the development of general guidance for the design and use of ATDs. As part of the analyses, ATD design and utilization questions were identified for which guidance could not currently be specified with a high degree of confidence.

Procedures for Initial Preparation of Research Topics

Based on their interpretation of the information collected or generated from the above sources, STRES team members were asked to record their ideas for potential research throughout the course of Phase II data collection and analysis activities. At the end of these activities, potential research topics were solicited from each member of the STRES team. Emphasis was placed on the generation of topics that would address critical Air Force ATD training problems and that would be feasible to conduct in the near future with existing research facilities.

The initial collection of potential research topics yielded some 130 items. To facilitate their review, the topics were organized according to the content of the proposed research (e.g., ATD design questions, training methods and techniques, instructor management, and performance measurement).

COMPARISON WITH EXISTING RESEARCH PLANS

Since it would be of little utility to develop detailed research plans for topics that duplicated or overlapped with research already planned by ATD research organizations, the potential research topics generated by STRES team members were compared to projects planned by agencies conducting ATD research within the topical focus of STRES.

Collection of Information Concerning Existing Research Plans

Agencies whose research plans were considered to be of interest included the Air Force Human Resources Laboratory, the Air Force Office of Scientific Research, the Naval Training Equipment Center, the Navy's Training Analysis and Evaluation Group, the Fort Rucker Field Unit of the Army Research Institute for the Behavioral and Social Sciences, the Army's Project Manager for Training Devices, the Federal Aviation Administration, and the National Aeronautics and Space Administration's Langley Research Center. Several of these agencies were visited by STRES team members during the Phase II data collection effort (see Appendix A). Documents were obtained, where available, that described current research projects and plans, such as five year plans and work unit summaries.[1] In addition, when necessary, personnel at these organizations were interviewed concerning their projects and plans in order to clarify the nature of their research. Since the Phase II effort was 29 months long, these personnel were also recontacted throughout the effort by telephone and mail to identify changes in plans that had occurred during Phase II.

---

[1]Formal five-year planning documents were available to the STRES team from only the Air Force and the Army Research Institute for the Behavioral and Social Sciences. Citations for the latest such documents relevant to the STRES effort are given below. For those agencies for which the project team had no formal plan documents, information was sought through discussions with research personnel and through review of less formal items such as work unit summaries, memoranda, and letters.

The formal documents used were:

[a]Five-Year Plan for R&D in Aircrew Performance. Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, 1978.

[b]Air Force Aircrew Training Devices Master Plan. Wright-Patterson AFB, OH: Deputy for Development Planning, Aeronautical Systems Division, March 1978.

[c]Flying Training Research Plan 1978-1982. Brooks AFB, TX: Air Force Human Resources Laboratory, March 1978.

[d]Letter, Commander AFHRL(FT) to AFHRL/XR, Subject, AFHRL/FT FY 81-85 Research and Technology Plan. 30 July 1979.

Procedures for Comparing Topics with Plans

To facilitate their coordinated review, each of the potential research topics generated by the STRES team was annotated with respect to the ongoing or planned projects, if any, that were similar to the research topic. If a topic appeared to overlap substantially with an already planned project, then the points of similarity and difference were also noted in the annotation.

This annotated topic list was then reviewed by the senior members of the STRES team. During this review, each topic was analyzed and discussed with respect to its similarity to existing plans or to other topics that had been suggested. Topics that overlapped or duplicated existing plans were eliminated from consideration for expansion into detailed research plans.

These topics were retained for purposes of discussion and inclusion in the present report. They appear in Appendix D, which lists all topics, whether developed into detailed research plans or not. However, potential research topics that were redundant with topics suggested by other STRES team members were reduced into a single topic that included all aspects of the research that was suggested originally.

The topics that survived this review served as input to the next activity.

PRELIMINARY ASSESSMENT OF PRIORITY

Given the impracticality of conducting research on all of these topics during Phase III of STRES, the potential research topics were classified with respect to their priority for research. This classification then served as a primary factor in the selection of topics as subjects for detailed research plans. The following paragraphs discuss the steps that were employed in the classification procedure.

Initial Review of the Topics

The list of remaining potential research topics, with information pertaining to the source of the topics, rationale, and expected impact, was circulated among selected members of the STRES team for their review. To provide a standard basis for review of the topics and also for communication between reviewers concerning the topics, the STRES team members were asked to rate the topics on three dimensions. These dimensions were: (1) the priority that the members felt the Air Force should give to the research topic based on the magnitude and criticality of the operational training problems that the research would address; (2) research feasibility in terms of the amenability of the problem to a research approach and also the extent to which operational constraints would restrict the research; and (3) delay of payoff for the research in terms of the amount of time expected to elapse before the results of the

research would impact on Air Force training programs. The listing of high value tasks and issues, previously mentioned, provided a background for this review of the topics. That information was integrated by the team members with their experiences over the course of the STRES project.

Analysis of the Initial Review

The STRES team members who reviewed the topics were then convened as a group to discuss the topics and their ratings of them. The group discussion was employed as a preliminary to the final classification of the topics. It allowed full exchange of information concerning the topics and elaboration of the rationale for their submission. As a result of the group discussion, certain of the topics were rewritten in ways that better reflected the group's consensus view of the topic, and some new topics were generated that the group felt better addressed significant problems. Largely, these revisions were matters of eliminating redundancies. This pruning and revision of the topics resulted in a listing of 111 topical items that served as input to the next step.

Classification of the Topics

Following the group discussion, the STRES team members were asked to review the list of 111 potential research topics, as modified in the previous step, and independently to classify the topics into one of three categories. These categories were: (1) a high priority category for those topics believed to warrant development into detailed research plans because of the magnitude of their potential impact on critical Air Force ATD training problems; (2) a secondary priority category for topics deemed to be significant issues for research, but of lesser impact than those in the first group; and (3) an exclusion category for those topics judged not to warrant research because they were either fundamentally not issues requiring experimental research (e.g., certain management issues), were not amenable to research, were issues for which additional research was not required, or were judged to be of little operational impact or concern to the Air Force.

Review of the Classification

The STRES team members who classified the topics were convened again for a group discussion of the classifications they had made. Each topic, along with its assigned classifications, was reviewed individually. When there was a diversity of opinion, the reasons for the divergence were discussed, and members were allowed to modify their classification as warranted by the discussion.

Following the discussion of all topics, a tally was made of the classifications assigned to each topic. This tally was then employed to place each topic into one of three groups. These groups were: (1) topics that had been recommended by at least one team member for inclusion in the high priority listing; (2) topics that were classified by

one or more members in the secondary category and that had not been classified by anyone in the high priority category; and (3) topics that had been classified by all participants in the exclusion category. These three groupings were then provided to the Air Force for preliminary review and comment.

COORDINATION OF TOPIC LISTINGS WITH THE AIR FORCE

To obtain an independent review of the topics and their assigned classifications, the list of 111 potential research topics, classified into the three groups discussed above, was submitted to representatives of the Air Force for review and comment (the full list of topics provided to the Air Force is presented in Appendix D). This review was accomplished in two phases. The first consisted of a briefing of AFHRL personnel by STRES team members at a meeting held at the Flying Training (FT) Division of the AFHRL (AFHRL/FT).[1] Each topic and the classification assigned it by the STRES team was reviewed individually during the briefing. Reasons for the classifications were presented, and the classifications were discussed by the assembled personnel in both supportive and exceptive terms, as appropriate. In the second phase, the topic listing was reviewed over a one month period, and written comments were then submitted through the AFHRL Contracting Officer's Technical Representative (COTR) to the STRES project team for integration as additional information.

It is worth noting at this point that these inputs were provided as informational and not directive in nature. It is virtually inevitable that there will be areas of disagreement across agencies and individuals in any assessment of the relative importance and criticality of research project needs such as in the present case. However, the Air Force took a firm position that the ultimate output of Step 5 of STRES should be a research plan that represented the integrated experience and views of the STRES project team in an independent manner. Air Force views expressed in these coordination activities, then, were viewed as additional data for consideration by the team, but not as directive guidance. Thus, the research plan described in this document does represent such an independent assessment by the STRES project team and, as such, may be at variance with some of the coordination comments received.

In turn, the Air Force will view the plans submitted here as an input for its consideration with reference to Phase III of the STRES project. Final priorities in that regard will be established by the Air Force and will be based on the relationships between the topics described in this report and the official research plans for AFHRL that exist now or will exist in the future. Official plans necessarily change as service needs and policies change, so priorities (as eventually established by the Air Force) for the topics presented in this report, may change also. For this reason, the rationale underlying each

[1]Now the Operations Training Division (AFHRL/OT).

of the topics and plans presented here is given so that the rationale can be examined for its continuing validity in light of future changes.

RESEARCH PLAN PREPARATION

The first part of this final Step 5 activity consisted of the selection of topics for expansion into detailed research plans. To provide a comprehensive basis for the selection process, a final listing of topics was prepared that included the following information for each topic: (1) the comments made by AFHRL either at the briefing or in written reply; (2) the comments by STRES Advisory Team members; (3) the classification of the topic into one of the three levels of priority by STRES team members; and (4) the number of STRES team members who had classified the topic originally into the highest priority grouping.

Based on this information, the topics were grouped into a final set of four categories. The categories employed are: (1) topics that are of critical importance to the improvement of current ATD training or to the development of future ATD training programs; (2) topics that are important with respect to their impact in ATD training, but which should be of a lesser priority than topics in the first group with respect to competition for scarce R&D funding; (3) topics that could provide information that would be potentially useful in the improvement of ATD training, but whose potential impact is significantly less than that possible with topics in the first two categories; and (4) topics that should not or could not be addressed through experimental research.

Topics in the first group became the subjects of detailed research plans. These plans, which are presented in Appendix B, include the following information: (1) statement of the problem; (2) suggested experimental methodology; (3) comments concerning types and numbers of subjects; (4) data collection plans and data analysis procedures; (5) facilities and recommended locations for the research; (6) projected schedules for accomplishment of the research; and (7) manpower requirements.

Topics in the second group are also the subjects of research plans, but the plans for these topics are of lesser detail than those for the first group. These plans are presented in Appendix C.

Topics in the third group are included in the Appendix D listing of research topics submitted to the Air Force for coordination purposes. Topics in the fourth group, the exclusion category, are also noted in Appendix D.

III. OVERVIEW

INTRODUCTION

The detailed research plans for some 21 topics judged of highest priority are contained in Appendix B. The topics are interrelated in a variety of ways and do not represent isolated research fragments with no integrative unity. Further, the program presented here should not be viewed in isolation from research activities already planned or programmed in the AFHRL/OT simulation research plan. Rather, the present program and the topical areas that comprise it represent those areas of incremental emphasis that the STRES team feels, based on its collective experiences, will provide a more comprehensive research and technology base for future Air Force simulator procurement and utilization programs.

A research program can be viewed as comprised of a series of projects that provide the necessary enabling mechanisms or knowledge base to achieve one or more general objectives. In the present case those general objectives relate to the eventual implementation of simulator training programs in the Air Force that are both training and cost effective. A major emphasis in those future simulator training programs will be the training of advanced combat skills, particularly in the continuation training (CT) context. More specifically, then, the goals of the simulator research program can be viewed as the provision of the technology base necessary to effective training support for the wide variety of operational aircraft the Air Force will employ in the 1980s and beyond. Some of the research projects presented here will support all or most aircraft systems (e.g., performance measurement research, visual systems research), while others may be more restricted in their application.

For purposes of discussion, the projects are classified into two major groupings--simulation technology and utilization technology. Within each of these groupings, several major research thrust areas are identified. The individual projects described in Appendix B are distributed over these major thrust areas. Table 1 provides an overview of the thrusts and projects in each of the two major technology areas. It should be noted that the order of listing in Table 1 is for convenience of exposition and does not imply anything concerning relative priorities among these topics.

TABLE 1. DETAILED RESEARCH PLAN TOPICS

I. UTILIZATION TECHNOLOGY

A. Instructional Process

1. Feedback and guidance in ATD instruction

B. Evaluation/Measurement and Instructional Process Control

1. Evaluation of student performance
2. Evaluation of crew performance

C. ATD Instructor Factors

1. Instructor training
2. Instructor evaluation

D. Locus of Control in ATD Instruction

1. Self-instruction in ATDs

E. Skill/Task Oriented Research

1. Advanced cognitive skills
2. Tactics development and dissemination
3. Crew training
4. Extended team training
5. Operational training programs
6. Operational tasks

II. SIMULATION TECHNOLOGY

A. Models for ATD Design

1. Model for predicting ATD training effectiveness
2. ATDs in proficiency evaluation

B. Visual Cues

1. Visual cue requirements
2. Scene content

TABLE 1 (Continued)

C. Motion-Force Cues

1. Motion-force cue requirements
2. Motion-force cueing mechanisms

D. Instructional Functions

1. Displays of the external visual scene
2. Instructor location
3. Remote display terminals

It should be noted in examining the projects listed in Table 1 that the overall emphasis is on utilization. Even those projects categorized as "Simulation Technology" have a strong underlay of concern with utilization. As has been noted previously, current AFHRL/OT research planning reflects a need for research targeting operational utilization needs, and the projects presented here, with their basic orientation toward learning process research, along with those described in current AFHRL plans, will allow significant enhancement of future Air Force simulator training. The present projects treat learning and instructional process factors that are critical to effective and efficient training.

The following discussion explains the rationale for the selection of the various topics. Also, some of the across-topic relationships are treated, as well as relationships between the topics and various aspects of Air Force simulation and training requirements.

UTILIZATION TECHNOLOGY

Observations made during STRES concerning current simulator utilization practices suggest that considerable improvement could result if simulator devices and training programs were developed on the basis of a <u>learning</u> task analysis rather than from an <u>operational</u> task analysis.[1] By this is meant that the emphasis should be on analysis of learning problems and how the ATD and its training program can best address those problems. Thus, the focus would be placed more heavily on cue-response discriminations and generalizations that comprise skills, and on

[1] Though not included here as a primary research topic, research to develop a learning task analysis methodology was suggested as a research topic (see Appendix C).

instructional process factors such as diagnosis, feedback, guidance, and mediation. These issues are common to all training situations, not just to the use of ATDs. Thus, there are broad benefits to be gained from casting future research in a more general context than one that is narrowly constricted to ATD questions. However, sight should not be lost of the unique aspects of ATD utilization problems.

It is of some interest to note that, while a fundamental aspect of the rationale for developing simulators is to allow facilitative management of the learning process variables without the restrictive constraints of the "real world," there often seems to be an almost complete absence of attention to such variables in the design of the ATDs themselves and of the programs for their utilization. The concern is usually with real world fidelity and the operational task, rather than with instructional facility and the learning task.

The ISD methodology, which, at least nominally, underlies the development of Air Force training media and programs,[1] provides an appropriate framework for directing attention to instructional process concerns and does provide for some learning analysis functions. However, it provides only general guidance concerning how to apply such analyses. General guidance that feedback, for example, is necessary to learning is not sufficient to the needs of either the device designer or of the device training program developer. What is needed is a specific information base covering factors such as how to provide feedback, when to provide it, how to handle the task specifics of feedback, and similar instructional management factors. As noted, it is in the facilitation of the manipulation and management of such learning process factors that the real essence and reason for simulation lie. Yet, in most devices and programs such factors are seemingly ignored.

The utilization research suggested here is aimed at extending certain aspects of the learning analysis methodology and other instructional factors as they relate to simulation. One major thrust, then, deals with instructional process factors. Another concerns the role of measurement and evaluation as factors in controlling or managing the instructional process. The ATD instructor represents another major variable in the quality of utilization, so there is a research thrust concerned with instructor-related factors. The locus of control of instruction

---

[1]It should be noted that in spite of the official acceptance of ISD by the Air Force for a number of years, none of the continuation training simulator programs observed in STRES had been developed on the basis of substantial ISD input. Further, the actual role of ISD or systematic learning analysis in determining the requirements for or design of ATDs must be viewed as minimal. In practice it has been difficult to initiate ISD efforts sufficiently in advance for them to influence training hardware design efforts.

(i.e., trainee, device, or instructor) is still another concern treated. Finally, though not so directly concerned with instructional processes per se, there is a thrust of research concerned with specific operational tasks, skills, or programs. This latter thrust is applications-oriented in a more specific manner than are the others mentioned. While these applications-oriented projects are not demonstrations, in the sense discussed previously with reference to technology transfer problems, they will have some demonstration value in addition to their research value. Their principal function is as prototype studies in programmatic research. The prototypical study may then be repeated for other operational tasks, as appropriate.

Instructional Process

One project is described in Appendix B dealing directly with instructional process variables. While there are numerous other aspects of instructional processes on which beneficial research could be conducted,[1] the one judged to have sufficient priority for inclusion is research on feedback and guidance as they affect simulator training of aircrew skills and supporting mediation processes. In addition, these factors are integral in many of the other research topics recommended. The potential instructional process topics not included were omitted for a variety of reasons. Some are treated in research that is already planned or are part of other research topics treated here. Some are nonexperimental in nature, while for others the existing knowledge base would seem sufficient for effective application.

The intent is that instructional process research be conducted in the context of analysis of the learning task, as discussed earlier, as opposed to the operational task. Obviously, the intent is not to ignore the operational task--it is the beginning point of analysis and of training requirements development--but, rather, to ensure extension of analysis beyond it to the learning task. A knowledge base on feedback, guidance, and mediation is essential if ATD design and ATD training programs are to be grounded in learning analysis. It is through manipulation of these process variables that simulation can make perhaps its most powerful contribution to effective training, a point that is both explicit and implicit in a number of the recommended research projects. Thus, research in the instructional process area is a key aspect of the overall program described here.

---

[1]Examples might include: part- vs. whole-task training; integration of separately learned tasks; effects of level of learning or overlearning; interaction of trainee factors with learning; adaptive training; and effects of alternative organization or sequencing of media/training tasks.

Evaluation/Measurement and Instructional Process Control

The general area of performance measurement and evaluation is of pervasive importance in training. Currently planned AFHRL/OT research reflects its criticality. One of the major technical goals of that program is the development of sensitive, accurate, and reliable quantitative measures of aircrew performance in order to provide means for evaluating aircrew skills, training methods, courses and devices. The results of the STRES effort underscore the need for such research. However, the results also suggest a need to extend the emphasis in measurement/evaluation research to include examination of the role of measurement as a controlling or pacing mechanism in the instructional process.

Consequently, the topics proposed in this thrust area deal with evaluation and instructional process control for individual training and for crew training situations. The emphasis on learning factors previously stated is basic to definition of these two research topics. The research proposed would treat the role of measurement in managing instruction (e.g., pacing, sequencing, etc.) and in the handling of factors such as feedback and guidance (e.g., when, what kind, how much, contingencies, etc.). In addition, the use of measurement for system quality control is also an aspect of this process management concern.

Unquestionably, a major factor in the instructor's management and delivery of instruction is his evaluation of trainee performance, whatever the basis for that evaluation. Research is required to identify more precisely those aspects to which the instructor should be reactive for instructional management and to develop techniques for their measurement/evaluation. While something is known of such factors with regard to individual training, relatively little information exists concerning such factors in crew training. The two projects proposed here will seek to provide a methodology and data base that will allow better instructional process control in ATD training.

ATD Instructor Factors

The training of ATD instructors is a significant factor in the effectiveness of any ATD program. The STRES observation that most ATD instructors function in terms of the "aircraft training model," along with the fact that ATD instructor training is often more oriented toward simulator operation than simulator use, suggest a clear need for more attention to the instructor as a training system process factor. In particular, improving the ability of the instructor to utilize fully and effectively the capabilities of simulation to manage instructional process functions is crucial to improving the effectiveness of ATD instruction. Consequently, instructor factors are an area of emphasis in the proposed program.

There has been a start on ATD instructor training research, but more work is required if the instructor's role is to be expanded to that of a real training system manager and dynamic interactive system element. Information needs, required skills and knowledge, student learning problem diagnostic skills, ability to use measurement indices for learning process control, and similar learning-oriented factors would be the areas of concern in such research.

Two research projects are described for this area in Appendix B. One concerns ATD instructor training--functional needs, role definition, training program design, evaluation of alternative training methods, etc.--while the other deals with the development of measures of instructor performance. Performance measures are required to serve as criteria in alternative methods evaluation studies, but they would also serve, or provide a basis for, another ATD system function need discussed in the STRES Utilization report, i.e., the evaluation of job performance of ATD instructors.

In a very real sense, the ATD instructor is the key effector mechanism in increasing the effectiveness of ATD programs. He is the principal interface between the utilization and simulation technologies and the learner. He can be an instructional force multiplier, or he can vitiate instructional effectiveness. The instructional outcome of ATD training is dependent in large degree on the manner in which the instructor subsystem is designed and implemented. Thus, this thrust is a critical concern.

Locus of Control in ATD Instruction

One consequence of the shift in emphasis of ATD research from UPT to the CT setting is a change in both the nature or level of skills involved and in the nature of the trainee. The UPT training environment generally provides carefully controlled individual learning experiences for the trainee under the control and scrutiny of an instructor. This is also true to a considerable extent of the Combat Crew Training (CCT) environment as well. As the pilot progresses up the skill and experience continua, there is a decreasing degree of direct instructor control of the learning experiences and a move toward combat-scenario oriented training. The locus and nature of instructional control shifts from instructor to trainee. While a considerable amount is known about instructor control of instruction, relatively little is known about the trainee as the locus of control.

The ATD presents still another dimension to locus of control in terms of the potential of the device itself as a controller. The considerable research that has been done on automated training and adaptive training in ATDs provides some data concerning this aspect of locus of control.

The matter of locus of instructional control assumes a real operational importance in terms of its relationship to the manner in which ATDs currently being procured will be used. During the various site visits made by the STRES team a substantial amount of self-instruction was observed in operational units. Since the bulk of the devices being procured are aimed at continuation training at operational units and a continuing shortage of instructor resources may be predicted, it may be expected that the amount of self-instructional use of those devices will be substantial.

In view of these considerations, the development of a systematic knowledge base concerning the efficacy of self-instruction as a function of task type, trainee type, level of skill, and similar factors is recommended. Such research will allow more effective employment of instructional resources in ATD programs.

Skill/Task Oriented Research

While the preceding utilization research projects are generally instructional-process oriented, another major thrust related to utilization deals with research on specific tasks or types of training. Such projects are related to a number of currently planned AFHRL/OT efforts concerned with operational training, particularly in the tactical area.

As has been noted before, the shift of ATD concern to the CT area has resulted in increasing attention to the training of higher order skills. One aspect of such is the cognitive skills area. The potential of ATDs for cognitive skills training is considerable, and a project is recommended to enable better exploitation of this potential as it relates to tactical requirements.

Another area of ATD potential that offers considerable promise is the use of ATDs in the development and dissemination of tactics. The emerging capabilities for free-play, two-sided air combat in ATDs, for simulation of the electronic environment of ECM and EW, and for simulation of various threats such as SAMs, all present a means for developing, evaluating, and disseminating tactics that may be more advantageous than airborne tactics development and dissemination. The ATD, in combination with ACMI facilities and appropriately designed flight scenarios, should allow significant gains in tactical skills and readiness if the research to exploit ATD capabilities in this area is carried out.

Projects are also suggested with reference to improved crew training in ATDs and extended team training. This latter term refers to the ATD training of pilots or crews in an extended interactive simulation context with other related personnel and systems (e.g., FACS, ground controllers, and ground force elements). The potential gains in effectiveness and cost savings from such system-interactive ATD training should be considerable if the research is done to support effective program development.

Extended team training might also be construed to include studies in which two or more simulation devices remote from one another might be linked together. While such efforts might prove fruitful in the future, the use here of the term "extended team training" is not intended necessarily to denote such integrative activities. The concern here is more with identifying extended team training needs of the full operational combat system and with developing means for meeting such needs through any simulation approach.

Finally, projects are suggested related to methodology for the development of full operational ATD training programs and for training methods studies related to specific types of operational tasks such as air-to-ground weapons delivery.

The projects suggested in this applications thrust area are similar in emphasis to the current direction of the planned AFHRL/OT ATD projects in operational training and support of the operational commands. The projects suggested here serve to underscore such emphasis as responsive to Air Force needs as perceived by the STRES team and will aid in securing the greatest return from the utilization of future Air Force simulation resources.

SIMULATION TECHNOLOGY

While the preceding discussion has highlighted the criticality of research needs in the utilization technology area, particularly with reference to instructional processes, there are also critical needs in the simulation technology area. Many of these needs have already been identified within the planned programs of AFHRL and other research agencies such as NTEC,[1] TAEG,[2] and ARI,[3] but additional effort is required. Some of the projects suggested here deal with content that is included in such programs, but with a different emphasis. Other projects suggested here represent additional thrusts or emphases not found in the agency programs. The topics developed here build on the plans of the various simulation research agencies, and together they represent a balanced approach.

Many of the current or planned agency efforts in the simulator technology and engineering areas are aimed at providing the basic hardware/software mechanisms necessary to increasing, or potentially increasing, training effectiveness and capability. Examples of such efforts include investigations of alternative force cueing systems, the development of voice recognition and synthesis systems, and the development of wide-angle visual systems. Additional visual technology efforts are aimed at areas such as edge capability increases and circle generation capability for CIG systems, improved image display quality, and improved data base generation techniques, all of which will provide benefits in the visual training area. Such engineering R&D is necessary

[1]Naval Training Equipment Center.

[2]Naval Training Analyses and Evaluation Group.

[3]Army Research Institute.

and encouraged to provide the mechanisms for implementing effective training in the future.

Research in these areas that is underway or already planned at AFHRL/OT, particularly in the visual technology area, reflects an increased emphasis on utility for operational training needs at the combat unit level. Examples of such operational emphasis include the programs related to A-10, F-16, and C-5 systems. In fact, operational command training program support and tactical combat aircrew research and development are two of the thrust areas already planned by AFHRL/OT for the FY 81-85 time period.

The simulation technology program presented in the present document contains projects in four major research areas. The emphasis in each of these four thrusts is on development of information to allow the design of more training effective simulator systems. The areas have some commonality with those already planned at AFHRL/OT in terms of simulation content, but they represent an additional and needed emphasis on the learning process understructure to system design.

Models for ATD Design

The first thrust presented involves two projects--one of broad conceptual significance, and the other of somewhat more specific scope. Both deal with development of improved guidance for ATD design, i.e., guidance that will make ATDs responsive to training and aircrew performance requirements in a more deliberate and informed way.

It is clear that the design of the ATD establishes certain limits or boundaries for the manner and extent to which the ATD training program can exploit principles that will enhance learning. While the need to extend our knowledge of the instructional process has been discussed in connection with utilization technology, there exists a companion need for better means of implementing that which is known into the details of ATD design in such fashion that training effectiveness is enhanced or maximized. In recognition of this need, a project is recommended here to develop and validate a model for relating training effectiveness to device design features. Such a model will have obvious utility for the development and evaluation of ATD design alternatives, and it would provide guidance for designers that was based in training effectiveness. In addition, though not its primary purpose, it could provide guidance for simulation research and could have utility for enhancing operational utilization of existing devices. The model would build on existing data as well as that to be generated in the research on instructional process variables previously described.

It might be argued that development of such a model would better be classified along with the utilization technology projects already discussed. However, it is included here because the focus is on development of a model for use primarily as a tool in the simulator design

process. A principal problem with the hardware and software designs of many ATDs is the apparent lack of any sound learning-based design guidance. The development of a model such as described here would provide such guidance.

The second project is concerned with design requirements related to ATD use for proficiency evaluation. Assessment of combat skill proficiency is a somewhat different matter than the assessment of training performance. As a consequence, there are differences in the functional requirements for such measurement, and these differences should be reflected in ATD design if ATDs are to be used for such purposes. This is not to suggest that there are no commonalities between ATD training-related assessment requirements and those for assessing operational proficiency; the commonalities are considerable. However, if ATDs are to be useful in providing operational combat proficiency assessments--and the practicalities of readiness assessment would support the desirability of using ATDs for such purpose--the functional requirements for such evaluation must be identified, system characteristics and design considerations developed, and validation studies conducted.

Visual Cues

Visual technology research is a major area of activity and research need. The increasing emphasis on tactical skills training brings with it a corresponding increase in concern with visual simulation technology. The AFHRL/OT program, using the ASPT and SAAC devices, is already embarked on an extensive visual technology research effort, one that will be intensified even more as a function of the increased attention being given to tactical task training. The two projects recommended here would extend the visual research already planned at AFHRL/OT into different aspects of visual scene information.

One project deals with an examination of visual cue requirements as a function of task factors. The critical aspect of representing the visual world is to represent the cues, i.e., the information necessary to learning and performance that is derived from a given set of visual stimuli, rather than to replicate the physical stimuli. The cue is the behaviorally critical factor, not the stimulus per se. Further, the stimulus-cue-behavior factors must be viewed in the context of their relation to transfer to the operational situation. Visual cue analysis and empirical validation research will provide a critical and necessary guide to visual technology and engineering development, as well as a guide to the effective use of visual systems.

The second project deals with visual scene content. There are obvious limits to scene content that are imposed by the visual technology used (e.g., CIG, camera-model, etc.). At the same time, there are scene content necessities that derive from the task to be taught (e.g., air-to-ground, low level navigation, etc.). Further, the scene content requirements and limitations must be viewed in terms of the cue

requirements previously discussed. A given cue (e.g., surface extension or depth) may be produced by several scene content alternatives (e.g., grid, relative size of familiar objects, surface texture). These content alternatives may carry quite different concomitant cue consequences. Therefore, content determination must be based on a broad and systematic look at task and cue structure.

Motion-Force Cues

A considerable body of research exists concerning motion-force cues and their utility for ATD instruction. However, there are divergent views as to the meaning of that research. One reason for this divergence is the relative emphasis in motion research on the stimulus (as opposed to the learning/performance cue orientation). Motion should be approached from the perspective of behavioral cue analysis rather than sensory stimulus replication. Further, the concern should be with the overall cue structure for a given behavior or task, e.g., questions related to the task specifics of disturbance motion vs. maneuver motion.

Another major aspect of motion-force cueing concern is that of alternative mechanisms for providing such cues. The provision of platform motion can be expensive, and it is not feasible to provide certain sustained motion-force stimuli through platform motion. As a consequence, there is considerable interest in alternative motion-force cueing mechanisms, particularly with reference to the use of simulation in air combat maneuvering tasks. Efforts are already underway concerning the development of alternative motion-force cueing systems.

Two projects are recommended here in the motion-force cueing area. They will provide a needed complement to current efforts and will reflect a task-behavioral emphasis rather than a stimulus emphasis. One project deals with determining motion-force cueing requirements for various skills, while the other deals with alternative mechanisms for providing such cues. The two efforts are companion pieces, and there should be a continuing cross-feed of information between them.

Instructional Functions

A major area of ATD design interest concerns those features of the devices aimed specifically at supporting the instructing function. While one might maintain that any feature of an ATD has that purpose, certain of the features have direct relevance to the provision of trainee feedback, monitoring and evaluation of performance, and to enhancing the ATD instructor's ability to carry out his functions. Such features may relate to either effectiveness or efficiency of the training operations. They are sometimes referred to as instructional support features (ISF). One of the STRES reports deals with ISF and the state of the art and practice.

It is in the ISF area that one would expect to see the clearest recognition of learning process considerations in ATD design. In fact, features such as "freeze" and "playback" can be viewed as deliberate departures from the fidelity orientation so prevalent in simulator design. As noted, the intent of such features is to allow deliberate intervention in and manipulation of the behavioral and learning processes. Unfortunately, however, the design and utilization of such features does not seem typically to reflect any systematic behavioral and learning analysis. In fact, many ISF, some of which are costly, go unused in practice due to inappropriate design and/or to lack of understanding of and training for their use by instructional personnel.

For most ATD training, the instructor is conceived as the controller of the training system in the overall sense, and as the manager of the training process and of the minute-to-minute flow of instructional events. Stated differently, he is a major aspect of the interface between the trainee and the instructional system. For these reasons, the design of the instructor/operator station (IOS), which would include most of the ISF, is a critical factor in the effectiveness and efficiency of the ATD as a training system. While IOS design was suggested by the STRES team as a research topic, it is not represented in Appendix B because of the existence of ongoing research projects on IOS design within the AFHRL structure, including the ISF work previously mentioned. However, the STRES team underscores the importance of IOS design research, because, to a considerable degree, it determines the nature and extent of learning process representation in ATD design and, ultimately, in ATD utilization.

Three specific projects are recommended in the instructional functions thrust portion of the present program. These, along with the instructional process research previously discussed, should provide important supplement to already planned IOS and ISF research within AFHRL. The first project deals with requirements and methods for displaying extra-cockpit visual scene information to the ATD instructor. The growing concern with visual simulation and training problems that has been described, along with projected future visual system procurements, makes this an area of critical design concern. Not only are there substantial practical and cost considerations, there is the fundamental consideration of what is required and desired with reference to the instructor's executing his visual instructional functions. Therefore, a project is recommended to determine the nature of what should be displayed and the instructional effectiveness of various techniques or systems for displaying such visual information.

The second project concerns the physical location of the instructor. Again, the design and cost impact of this question is considerable. The location of the instructor determines many aspects of his manner of interacting with the trainee, and it also determines many aspects of the information that can be provided the instructor. This matter appears of considerable concern with reference to ATDs designed for

air-to-air combat training. This research should be closely coordinated with that just described concerning display of visual scene information to the instructor.

Remote display terminals for post-flight feedback and debriefing purposes are the subject of the third project. While most existing ATDs have at least some means for post-flight feedback (e.g., hardcopy traces and print-outs), their ability to recreate the dynamics of the training session is usually limited. Further, dynamic replay makes the ATD unavailable for use by another trainee during the replay period.

The literature on human learning, as well as experience with specific training and evaluation systems such as ACMI, would indicate that post-flight feedback should be an extremely significant part of the training value of ATDs, particularly in combat scenario training such as air-to-air encounters. As the capability of simulators to provide such complex training increases, the question of providing post-flight feedback becomes more critical. Consequently, a project is recommended to examine the utility, functional requirements, and design implications for remote display terminals.

COMMENTARY

The projects in Appendix B comprise the research program recommended by the STRES team. Those projects share the common orientation of providing, in conjunction with projects already planned, the technology base to allow significant increases in the effectiveness and efficiency of future Air Force ATD training. It is likely that available resources will not permit the conduct of all the projects, even though they are all judged to be of high priority. It is recognized that research program thrusts change over time just as do operational needs. Thus, the program recommended here, which represents a "snapshot picture" of research needs as perceived at this time by the STRES team, should be examined by the Air Force and modified as indicated by different perspectives and changes in events. However, it is felt that the thrusts and emphases of this program represent critical areas of need that must be addressed.

The shift in focus of simulator research and applications to the operational combat skills areas is reflective of the growing role of simulation in Air Force training and readiness plans. While the state of the art in both simulation and utilization technologies is relatively advanced with reference to areas such as UPT, transition training, instrument training, and procedures training (though there is room yet for significant progress), the shift in emphasis to the operational skills area brings many significant challenges to those technologies. Correspondingly, there are engineering challenges that derive from this shift, and they will require R&D support. Further, there are aeromedical questions that warrant investigation in relation to training simulation (e.g., the post-flight psychophysiological effects of SAAC training

sessions). Thus, full realization of the potential future benefits from ATD usage in the Air Force will require a broad coordinated research effort. The program recommended here would be a significant part of such effort.

APPENDIX A

SITES VISITED BY STRES TEAM

PRECEDING PAGE BLANK-NOT FILMED

The various sites visited by the STRES team are listed in Tables A-1 and A-2. Table A-1 lists those sites where visits were made to ATD training facilities and organizations. The sites at which visits were made to agencies and organizations not actively involved in the direct conduct of ATD training are listed in Table A-2. The latter visits were generally to agencies involved in training device research, procurement, test, management, support, or manufacturing activities.

TABLE A-1. TRAINING SITES INCLUDED IN PROGRAM SURVEYS

Sites and Units	Topics of Interest
Altus AFB, OK (MAC) 443rd Military Airlift Wing	C-5 transition training
Castle AFB, CA (GAC) 93rd Bomb Wing	KC-135/B-52 transition training
Denver, CO United Airlines Flight Training Center	DC-10/8-737/B-747 transition and continuation training
Eglin AFB, FL 33rd Tactical Fighter Wing	F-4 continuation training
Fort Rucker, AL U.S. Army Aviation Center	UH-1 and CH-47 undergraduate and transition training
Langley AFB, VA (TAC) 1st Tactical Fighter Wing	F-15 continuation training
Mobile, AL U.S. Coast Guard Aviation Training Center	HH-3/HH-52 transition and continuation training
NAS, Cecil Field, FL VP-174 and Light Attack Air Wing One	A-7E transition and continuation training
NAS, Jacksonville, FL VP-38 and Patrol Wing Eleven	P-3C transition and continuation training
Plattsburgh AFB, NY (SAC) 380th Bomb Wing	FB-111 transition training
Reese AFB, TX (ATC) 64th Flying Training Wing	T-37/T-38 undergraduate pilot training
Tinker AFB, OK (TAC) 552nd Airborne Warning and Control Wing	E-3A transition and continuation training

TABLE A-2. SITES VISITED FOR MANAGEMENT, RESEARCH, DEVELOPMENT, ENGINEERING AND COST SURVEYS

Sites and Agencies	Topics of Interest
Washington, DC Pentagon Headquarters, USAF	Management of Air Force ATD resources, and life cycle costs
Randolph AFB, TX Headquarters, ATC	Management of the use of ATDs in undergraduate pilot training, and life cycle costs
Langley AFB, VA Headquarters, TAC	Management of the use of ATDs in fighter aircrew training, development of ATD requirements, and life cycle costs
Eglin AFB, FL (TAC) Tactical Air Warfare Center	Procurement, development and evaluation of ATDs
Luke AFB, AZ (TAC) 4444th Operational Training Development Squadron	Development of training and ATD requirements
Williams AFB, AZ Air Force Human Resources Laboratory (AFHRL/FT)	ATD research
Wright-Patterson AFB, OH Air Force Human Resources Laboratory (AFHRL/AS)	ATD research
Fort Rucker, AL US Army Research Institute for the Behavioral and Social Sciences	ATD research
Langley, VA NASA Langley Research Center	ATD research
St. Louis, MO McDonnel Douglas Corp.	ATD design and research
Binghamton, NY Singer-Link Corp.	ATD design, procurement and evaluation
Orlando, FL Navy Training Analysis and Evaluation Group	ATD research and life cycle costs

TABLE A-2. Continued

Sites and Agencies	Topics of Interest
Orlando, FL Naval Training Equipment Center	ATD research and life cycle costs
San Diego, CA Navy Personnel Research and Development Center	ATD research and life cycle costs
Orlando, FL US Army Project Manager for Training Devices (PM TRADE)	ATD research and life cycle costs
Hill AFB, UT (AFLC)	ATD life cycle costs
Hollomon AFB, NM (AFTEC)	ATD life cycle costs
Luke AFB, AZ (TAC)	ATD life cycle costs
Offutt AFB, NE (SAC)	ATD life cycle costs
Scott AFB, IL (MAC)	ATD life cycle costs
Travis AFB, CA (MAC)	ATD life cycle costs
Williams AFB, AZ (ATD)	ATD life cycle costs
Wright-Patterson AFB, OH (ASD)	ATD engineering and life cycle costs

APPENDIX B

DETAILED RESEARCH PLANS

INTRODUCTION

The topics listed in this appendix comprise the research program recommended by the STRES team. The detailed research plans that are provided for each topic contain the following information: (1) statement of the problem; (2) research overview; (3) analytic requirements; (4) experimental method (this section is divided into two parts that deal with considerations for experimental control, and procedures); (5) subjects; (6) data collection and analysis plans; (7) facilities; and (8) schedule and contractor personnel requirements.

Comments Concerning the Research Plans

The general structure of the plans that follow is based on the eight informational subsections stated. However, to provide a general overview of the manner in which several factors are treated in the various plans, the following discussion is offered.

Measuring instruments. All of the experimental designs require administration of measuring instruments. The validity of the results of each empirical effort depends on the adequacy of these measures. At present, adequate measures, by and large, have not been identified. For this reason, the pre-empirical analytic efforts for most of the studies include the development of appropriate measures, and it is not possible to specify the nature of the measures until the analytic effort is completed.

As a result, adequate measures must be identified on an ad hoc basis for the particular tasks involved in each study. Because of the scope of each effort, it is not possible at this time to anticipate precisely the appropriate measures to use. Whatever the measures may be, however, they should have demonstrable reliability and validity, and be usable under the conditions characteristic of the study involved. Reliability is generally enhanced as the objectivity of measures increases, and to the extent standard conditions for obtaining them are achieved. Validity--ensuring that measures obtained reflect the information desired--is the crucial issue, however, and in a number of instances it may be necessary to fall back on expert judgmental evaluations of performance. In such cases, preliminary tryouts of the measures should establish inter-rater reliabilities greater than .75 and conditions, criteria, and procedures for deriving judgments should be adapted as necessary until acceptable reliabilities are obtained.

Numbers of subjects. Minimum numbers of subjects (or degrees of freedom) are identified for all empirical studies. These numbers are arbitrary to some extent, but it generally would not be wise to decrease them. The following is a brief description of what is involved in deciding how many subjects are needed.

First, one essential question concerns "degrees of freedom" (df). Although a technical concept, df can be handled simply and directly for present purposes: the larger the df, the less chance that sampling fluctuations will confuse interpretations of group differences, or differences between observations on a single group obtained under two or more experimental conditions.

A second essential question in deciding the number of subjects needed in a study concerns the magnitude of the experimental effect, i.e., group differences, relative to the differences among subjects within groups. If groups are very homogeneous, if one subject in a group is highly similar to any other in the group insofar as the behavior of interest is concerned, then small group differences can be statistically significant. If subjects within groups are quite variable, even substantial group differences may not be distinguishable unless a large number of df are available.

It is in view of the inter-related issues of possible group differences, similarity of subjects, and available df, that all recommended Ns and df are somewhat arbitrary. But as pointed out, it would not be wise to decrease them.

In a few proposed studies, correlational analyses are recommended. In these cases, stability of statistical data describing each group is the central concern, so Ns generally must be substantially larger than those needed for comparisons of group means. In these cases, the larger the N the better, because the purpose is not to identify a difference of practical importance, but a statistical description characteristic of a group. Correlation coefficients are notoriously unstable from sample to sample unless Ns are substantial. Even Ns of 50 should be considered quite small. If multiple regression analyses are involved, Ns would have to be on the order of 1,000 if only one study is to be run. One way around this problem of large Ns is to cross-validate findings in one experiment by repeating the study with a new sample. Such a procedure is recommended in one proposed research design.

Facility requirements. Each of the research plans has a discussion of the facilities that would be required to conduct the research described. In some cases, specific research devices are mentioned, such as ASPT, SAAC, and the like, while in other cases reference is made to use of existing operational devices that support current UPT, CCT, or CT training programs.

For each study requiring use of a simulation device or facility, an estimate is provided of the number of hours of device time required to support the research as outlined. These projected facility requirements are stated in terms of the tasks within the plan that they support, and in terms of the calendar time over which the device time will be distributed, e.g., 100 hours over a period of six months. The intent of these

facility requirements forecasts is to provide Air Force resource managers with information that will allow an assessment of facility scheduling and availability impact for these plans when considered along with other Air Force plans. It should be noted, however, that in many of the instances involving operational training programs, the simulator time requirements shown here would not be in addition to existing training time loads, but would substitute for or tradeoff against already programmed time.

Some of the plans presented here might require that a given device be reconfigured or modified to support the research. Such modifications may require engineering support and time over and above that cited for the research itself. Since the future configuration of a particular device cannot be foreseen at this point (e.g., the ASPT might have a variety of possible specific cockpit configurations), no time allowance is forecast for engineering changes in the various plans presented. However, such needs must be considered at the time specific action is initiated to implement any of the present project plans.

Schedule and personnel requirements. The final section of each plan presents a projected schedule of calendar time to conduct the research described and an estimate of the level of contractor professional effort (in person-years) required. These projections are given separately for the various tasks that are described for a given research project. For each such task, the schedule shows a start and finish time in terms of months after contract (MAC). Start times can be interpreted as meaning the beginning of the MAC indicated, whereas finish times refer to the end of the MAC indicated. For example, if the schedule shows Task III of a given project as starting 9 MAC and finishing 18 MAC, this denotes a start early in the 9th MAC and a finish toward the end of the 18th MAC.

RESEARCH TOPIC LISTING

The topics listed below are in this appendix. Page numbers are included to facilitate locating each plan.

1. TECHNIQUES FOR PROVIDING FEEDBACK AND GUIDANCE

PROBLEM

The adequacy of feedback and guidance provided during practice is a major determinant of training efficiency. A distinct advantage of ATD training is that instructors have more opportunities and, because of instructional support features (ISF), often more effective ways to provide feedback and guidance than they would have in aircraft. However, instructors in many ATD programs do not fully exploit either the opportunities or the ISF. Further, it often appears that the design of ATDs and their features reflects less knowledge of the role of feedback and guidance in the training process than would be desired.

The problem is that the considerable body of knowledge concerning feedback and guidance has not been formulated as applied principles that could be used discriminatively by nonspecialists in learning technology. The formulations that have been attempted were derived from specific learning contexts, usually laboratories. The resulting principles are often inconsistent, contradictory, or simply inapplicable in the varied context of training.

Thus, ATD instructors must depend on their own experience for intuitive guides in employing feedback and guidance. Intuition is often a poor guide; and when it is insightful it is still lacking in the comprehensiveness and specificity that well formulated applied principles for using feedback and guidance could provide.

Optimum use of ATDs in training will require that these two most basic contributions that can be made by instructors be understood by them. The first step is to formulate the principles governing use of feedback and guidance in a way that ATD instructors with a practical minimum of training can understand. Then, not only would they be more alert to opportunities and contingencies for employing these influences, but better use--and probably improved future designs--of ISF would result.

The purpose of the research proposed here is (1) to formulate principles for using feedback and guidance in ATD instruction; (2) to test the adequacy of the formulations by how well instructors can follow them; and (3) to determine empirically the adequacy of the principles through the effects of their use on student learning. The original effort would focus on one particular set of subtasks to be trained, preferably those involved in a high-value task such as approach and landing. A decision matrix would be constructed identifying cue-response discriminations to be made, options for using feedback and guidance, and contingencies for their use. The ATD targeted by the analysis, and used in the empirical test, should have a variety of state-of-the-art ISF.

RESEARCH OVERVIEW

The study would consist of the five major tasks shown below.

Task 1: Construction of a decision matrix identifying cue-response discriminations to be made, options for using feedback and guidance, and contingencies for their use; development of a scenario for testing instructors, use of feedback and guidance.

Task 2: Development and conduct of an experimental ATD instructor training program designed to teach the principles and use of the feedback/guidance decision matrix.

Task 3: Use of a test scenario to compare the instructional performance of instructors who receive the experimental training with that of instructors who do not.

Task 4: Assessment of the differences between the performance of students who receive training from ATD instructors employing the decision matrix and that of students who are taught by ATD instructors using conventional techniques.

Task 5: Final data analysis and preparation of the final report.

ANALYTIC REQUIREMENTS

Three types of efforts would be pursued at the outset: (1) analyses of subtasks involved in overall task performance in terms of required cue and response discriminations; (2) compilation of principles governing the use of feedback and guidance as presently formulated (some excellent literature sources for "abstract" principles now exist); and (3) analyses of contingencies for using feedback and guidance that relate to the status of cue-response discriminations by the student. The first effort would go beyond usual task analyses that focus only on correct cues for exercising a response, and the nature of the correct response. It would be necessary to specify how cue recognition can be learned and to anticipate sources of difficulty (e.g., interference from similar cues, deficient subordinate skills such as cross-checking instruments rapidly) so that possible bases for inabilities to discriminate could be specified. Similarly, the nature of evolution of responses from typical crude beginnings to finely tuned skill patterns should be understood so that an instructor could know not only what an action should be, but how a student progresses to skill perfection. It is through such knowledge that the learning principles referred to in the second effort can be made operationally meaningful, and that the contingencies for applying principles can be identified. For example, it is necessary to know whether or not a student's skills are evolving in an optimum way for learning, and if not, the nature of the desirable and undesirable

aspects. Positive and negative feedback could then be provided according to specific needs, and guidance could focus on the skill elements that go astray.

Either an existing or a newly prepared training regimen for teaching the skills involved would then be analyzed, so that a decision matrix could be constructed to apply across all component skills, but with branching as needed for separate skills. (The branching would be of special benefit for optimum ISF use.) If it is necessary for experimental and control groups to be at different training sites, the skills selected would be such that they would normally be taught at all sites used.

EXPERIMENTAL METHOD

Considerations for Experimental Control

Experimental controls must be adequate to allow answers to two questions: (1) Can the principles, as extrapolated in the decision matrix, be understood by instructors with a practical minimum of training? (2) Does use of the principles, as formulated, enhance skill acquisition in the simulator and, if feasible to test, transfer of ATD skills to aircraft? For the first question, comparable groups of instructors with and without training in the principles would be tested in a standardized scenario depicting various stages and problems of student learning. For the second question, a standardized training regimen would be followed by comparable groups of students, taught by comparable groups of instructors except for prior training on the principles and use of the decision matrix. Care would be taken that instructors in one group did not interact with those in the other. All personnel involved in implementing the experiment would be trained to fulfill their roles in a standardized manner.

Procedure

If feasible, subjects, whether instructors or students, would be assigned randomly to groups. If conditions require that experimental and control groups be at different training sites, baseline data would be collected on any variables likely to affect the dependent variables. For the first test, one group of instructors would be trained for their responsibility in the usual manner. A second group would be trained according to a plan developed during the experiment that was designed to teach the principles and use of the decision matrix. Both groups would be tested using a standard scenario as above. At standardized points in the scenario, all subjects would be queried as to needed instructor interventions (i.e., provision of feedback and guidance) and reasons for the actions selected. Measures revealing functional knowledge and use of principles of providing feedback and guidance would be obtained at each decision point. In addition, instructors would be encouraged in a

standardized manner to identify occasions for intervention on their own, and measures would indicate adequacy of recognition of problems as well as functional knowledge. (Measures would have to allow for alternative proper uses of principles.)

For the second test, separate groups of aircrew trainee subjects with usual variations in experience levels would be taught using a standardized regimen. Some of the student groups would be taught by instructors with usual or standard instructor preparation, while other groups would be taught by instructors who had the special training and who had demonstrated a good functional knowledge of the principles and use of the decision matrix.

All students would transition to an aircraft following ATD training.

SUBJECTS

All instructors (and students) used as subjects should be drawn from pools typical of those providing (or undergoing) the training. Twenty instructors should be used for the first experimental test, 10 in each group. Data revealing rate and level of achievement would be collected for at least 3 students taught by each of at least 5 instructors from each original group.

DATA COLLECTION AND ANALYSIS

Measures revealing instructors' functional knowledge and use of principles of feedback and guidance would be obtained as indicated earlier. Measures of student proficiency on separate skills would be obtained so as to reveal rate of learning in the ATDs and level of achievement at the end of device training. If feasible, measures would also be obtained for determining original level of proficiency when transitioning to the aircraft, and rate of progress to full proficiency in the aircraft. Analyses of variance would compare instructor groups, and measures on students, separately by experience levels.

FACILITIES

The ATD employed in this study should have a variety of state-of-the-art ISF. The research could be conducted using an operational ATD such as the UPT-IFS or a laboratory simulator such as the ASPT. The simulator that is selected for this study would be employed for a total of approximately 395 hours. It would be employed in Task 1 for a total of approximately 15 hours distributed over 4 months, in Task 2 for a total of approximately 100 hours distributed over 3 months, in Task 3 for a total of approximately 40 hours distributed over 6 months, and in Task 4 for a total of approximately 240 hours distributed over 11 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

Approximately 22 calendar months and 3.50 person-years of professional contract labor would be required for the study. The calendar time and level of effort required for each task and the total study are as follows:

Task	Schedule (Contract Months) Task start	Task finish	Number of person-years
1. Construction of the decision matrix and development of test scenarios.	1	6	1.00
2. Development and conduct of instructor training.	5	7	0.20
3. Evaluation of instructor performance using test scenarios.	8	13	0.50
4. Evaluation of instructor instructor performance using student performance.	8	18	1.00
5. Data analysis and report preparation.	18	22	0.80
	Total person-years		3.50

2. ASSESSMENT OF TRAINEE PERFORMANCE

PROBLEM

Assessments of trainee performance are needed to evaluate trainees and training programs. In addition, assessments that pinpoint the nature of trainees' learning problems are necessary to adapt training practices to the correction of problems, i.e., to manage and control the learning process. Because ATD instructional support features and the conditions of ATD practice can be designed to optimize these learning process control adaptations, it is necessary that reliable, valid, and usable assessments be available to instructors and trainees when needed if ATD training efficiency is to be maximized.

Over the past three decades, considerable effort has been made to develop measures and assessment techniques for evaluating aircrew, especially pilot, performance. However, almost all of the measures until recently focused on aircraft performance. Furthermore, the emphasis in the past, and now, has been primarily on developing objective measures so that their reliability (consistency) could be improved. Relatively little attention has been given the more critical problem of validity: Do the measures tell us what we need to know?

The failure to identify a nucleus of objective, valid measures that can be used readily in the control of ATD instruction and in ATD research is evident from a variety of current practices. At one extreme, root mean squared error is often the only objective measure used. The information it provides typically lacks the specificity of detail and timeliness needed to guide ATD instruction in process. Also, it is clearly irrelevant to many assessment needs.

At the other extreme, some research projects dealing with ATD learning have used up to 500 different measures, and 100-200 are common in some quarters. Various researchers' satisfaction with their findings notwithstanding, practically nothing has really been learned from such efforts, because the multivariate techniques of analysis were often applied inappropriately, and they did not begin to answer the basic question: Which, if any, of these measures tell us what we need to know?

Recent efforts in automated performance measurement in ATDs and in aircraft point to a possible significant breakthrough in measurement technology. We are learning promising new techniques of measuring. We should now find out what to measure for what purposes.

The goal of the research outlined below is to identify a nucleus of measures that are valid for defined purposes and usable in directing ATD instruction. It is recognized that validity is a highly specific concept, that a measure valid for one purpose, such as pinpointing a discrimination learning problem, may be useless for another, such as

evaluating proficiency. Also, a measure valid for diagnosing difficulties in performing one task may be inapplicable or even misleading when used for another task.

RESEARCH OVERVIEW

The study would consist of the three major tasks shown below.

Task 1: Selection of tasks for study and development of a candidate set of measures of trainee performance.

Task 2: Conduct of ATD training and collection of data using the candidate set of measures.

Task 3: Analysis of the data to determine the validity and utility of the measures, and preparation of the final report.

ANALYTIC REQUIREMENTS

The analytic effort would have two major thrusts: (1) to identify measures likely to reveal the quality of performance; and (2) to identify measures of value in pinpointing training problems so that timely constructive intervention during training can occur. For both thrusts, existing automated performance measurement technology would be surveyed for possible means for instrumenting measures, but the analyses of measurement requirements and development of techniques should not be limited by existing technology. Extensive and intensive reviews of the literature would be completed, and a set of tasks would be identified as a prototype for developing candidate measures. Instructors and students involved in training these tasks would be interviewed to identify training problems and assessment needs. However, neither instructors nor students generally would be able to conceive of the problems and needs in terms of basic discrimination learning processes, so a learning analysis would be completed to relate problems and instructors' informational needs to underlying issues of discrimination and generalization.

The learning analysis would go beyond usual task analyses that focus only on correct cues for exercising a response, and the nature of the correct response. It would be necessary to specify how cue recognition can be learned and to anticipate sources of difficulty (e.g., interference from similar cues, deficient subordinate skills such as cross-checking instruments rapidly) so that possible bases for inabilities to discriminate could be specified. Similarly, the nature of evolution of responses from typical crude beginnings to finely tuned skill patterns should be understood so that an instructor could know not only what an action should be, but how a student progresses to skill perfection. It is through such knowledge that the contingencies for instructor

intervention can be identified, and thus the types of information needed in a measure. For example, an instructor should be able to determine whether or not a student's skills are evolving in an optimum way for learning, and if not, the nature of the desirable and undesirable aspects. He could then adapt positive and negative feedback to specific needs, and guidance could focus on the skill elements that measures have revealed to go astray.

Analyses of this sort would provide an a priori basis for validating measures that pinpoint training problems. The issue would be simply, was the measure sensitive to indications of a learning problem, and specific enough to identify the underlying cue, processing, or response discrimination difficulty? Validation of proficiency measures would rely heavily on expert judgment as explained below.

The product of the analytic effort would be a list of candidate measures, techniques for obtaining them, and likely purposes that they could fulfill, particularly in instructional process control.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The empirical portion of the research would focus on the validation of the measures and on their usability for assessing students and directing instruction. In either case, measures should be obtained through standardized procedures that are feasible operationally. "Standardized" would have to be a somewhat inclusive description of procedures for obtaining measures to guide moment-to-moment instructional decisions, for if not automated, they must be applicable when and as needed. The task would be to define classes of procedures and conditions within which validities would be more or less invariant.

ATD task performances of aircrew personnel with a wide range of proficiency in the tasks would be used to establish validity for proficiency assessment. Validity or usability for diagnostic purposes would be determined in a separate effort utilizing the measures during standard training scenarios with typical trainee personnel serving as subjects. Instructor personnel would be trained to obtain the measures as specified, including criterion data. Instructors would be trained how to use measures for instructional decisions. They would also have to be proficient in principles governing use of feedback and guidance.

Procedure

For the variable proficiency group of subjects, tasks would be performed in the ATD under standardized conditions, and measures would be obtained as specified. In addition, criterion data would be collected to establish quality of overall performance and proficiency on the various

tasks. (The criterion would generally be subjective evaluations, because at present expert judgment is the best basis for determining aircrew proficiency. Enough independent judges would be used, however, to assure high reliability of average evaluations.)

The training group would undergo a standardized training regimen to learn the tasks. Measures would be obtained as specified in a standardized format, with additional measures instigated by the instructor if desired for decisions regarding his intervention during practice.

SUBJECTS

At least 100 aircrew trainee subjects would be needed in the variable proficiency group used to validate the measures for proficiency assessment, and at least 12 in the training group during the instructional utility study. The former group should vary, as stated above, in proficiency on the tasks used. At a facility where students are at different levels at any one time, all students could be tested for proficiency purposes during the same period. The training group should be typical of the range of students normally beginning training on the tasks, and they should be tested for diagnostic purposes throughout their training. Instructors in the latter case would be those normally assigned ATD instructing duties.

COLLECTION AND ANALYSES OF DATA

The variable proficiency subjects would be randomly divided into two groups of equal size. Canonical correlational analyses would be run for each group, adapted as needed to identify interrelations of clusters of predictor and criterion measures. Optimum, relatively small groups of predictors would be formulated as regression equations to predict particular criterion values and clusters for each subgroup. Each equation would then be cross-validated on the alternate group to identify the stable predictors.

Instructors of the student group would be queried via questionnaires and interviews as to the value and feasibility of measures for guiding instructional decisions. Students would be queried also, focusing on their perceptions of training problems and the value of instructors' interventions. These data would be collated so as to guide needed revisions in the contents of the measures and procedures for obtaining them.

FACILITIES

Facility requirements would be driven by the need for an automated performance measurement system that can be employed to obtain the

selected measures. A laboratory ATD such as the ASPT or SAAC should be used due to the probable need for engineering development of the measurement system. The simulator that is selected for this study would be employed for a total of approximately 235 hours. It would be employed in Task 1 for a total of approximately 15 hours distributed over 6 months, and in Task 2 for a total of approximately 220 hours distributed over 8 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

If an existing automated performance system can be modified and employed for the study, approximately 20 calendar months and 3.00 person-years of professional contract labor would be required. The calendar time and level of effort required for each task and the total study are as follows:

	Schedule (Contract Months)		
Task	Task start	Task finish	Number of person-years
1. Selection of tasks and development of measures.	1	8	1.30
2. Conduct of training and collection of data.	9	16	0.70
3. Data analysis and report preparation.	15	20	1.00
	Total person-years		3.00

3. ASSESSMENT OF CREW PERFORMANCE

PROBLEM

The effectiveness and efficiency of overall performance as an aircrew depends upon more than the competence of individual crew members in their separate jobs. They must do their jobs well, but they must also function as a team. Actions must be timed and coordinated, often precisely. Responsibilities must be assumed as situations and conditions require. Multiple supporting roles must be filled, including substitute roles. Communications among crew members must be clear, accurate, and adequate for the need. Withal, harmonious relations among the crew should be maintained.

The use of ATDs for team training will steadily increase. To exploit the specialized instructional opportunities ATDs can provide, team process measures, evaluative and diagnostic, will be necessary so that instructional strategies and practices can be evaluated and corrective training prescribed. Assessments of crew performance at present are based almost entirely on subjective judgments. A reliable, objective, demonstrably valid, and usable method for measuring team performance is needed. The measure would allow for the prerequisite condition of individual competence, while describing team performance as a whole. It would also permit diagnostic analyses to pinpoint areas of deficiencies to be corrected through training.

The primary purpose of the research outlined here would be (1) to develop methods for measuring overall aircrew performance that would provide the diagnostic and evaluative information needed for directing and controlling crew instruction in ATDs; and (2) to validate the measures empirically. In the process of developing and validating these measures, a prototype scheme for evaluating crew proficiency would be developed as well.

RESEARCH OVERVIEW

The study would consist of the three major tasks shown below.

Task 1: Selection of tasks for study and development of a candidate set of measures of crew performance.

Task 2: Conduct of ATD training and collection of data using the candidate set of measures.

Task 3: Analysis of the data to determine the validity and utility of the data, and preparation of the final report.

ANALYTIC REQUIREMENTS

The primary analytic tasks are (1) to identify the dimensions or parameters of aircrew performance; (2) to analyze team performance in terms of learning processes; (3) to identify types of information needed to direct and control the learning; and (4) to develop measures to provide the information when needed. Literature reviews and intensive interviews of aircrews would be necessary to define the scope, needs, and peculiarities of various aircrew constituents. In all likelihood, aircrew team performance can best be described as dynamic systems. Hence, the literature review would include in-depth analyses of systems functioning that depend heavily on personnel roles. These systems analyses would not only aid in identifying formulations of team dimensions and parameters, but would permit capitalizing on measurement methodologies that have been developed for various systems.

A preliminary measurement scheme would be developed for assessments of an aircrew with a fairly wide range of teamwork and instructional requirements so that the adaptability of the scheme to assessing a variety of types of aircrews could be assured. An ATD program for training crews to be used in the study would also be developed. Here, too, existing literature on training as related to systems requirements would probably be very helpful, for training must focus on the interactive aspects of roles, not just on the peculiarities of individual responsibilities.

EXPERIMENTAL METHOD

Considerations for Experimental Control

Experimental controls would have to be adequate for two types of validation: (1) instructor assessments of adequacy of measures for instructional decisions and evaluations; and (2) improvement in measures as crews improve during training. The first type of validation would be the primary concern, but the second would identify measures that reflect crew proficiency, and hence that would have potential in aircrew evaluations. All personnel involved in implementing the study and obtaining measures would be trained to fulfill their roles in a standardized manner.

Procedure

Aircrews would be trained according to the program developed during the analytic effort. Measures of aircrew performance would be obtained as specified by the scheme developed during the analytic effort. Measures would be obtained periodically beginning with the earliest stages of training as a crew and continuing through the end of training. Through questionnaires and standardized interviews, instructors would assess the value of the information obtained for instructional decisions and evaluations.

SUBJECTS

Subjects would be from the same type of aircrew, but varying in proficiency as required for the several types of data. At least 8 student aircrews should be used, and at least 4 instructors.

DATA COLLECTION AND ANALYSIS

Data would be collected as specified, either routinely or at the initiation of the instructor. Plots of measures over time for the crews used in the second type of validation would reveal how individual measures changed with improved performance. Instructor assessments would be collated and summarized. Information obtained from each type of effort would be used to revise the measurement scheme so as to increase validity, reliability, and usefulness.

FACILITIES

The study would be conducted in a CCT or CT unit in which crew training is conducted and the training resources of the unit would be employed in the investigation. The ATD would be employed for a total of approximately 335 hours. It would be used in Task 1 for a total of approximately 15 hours distributed over 6 months, and in Task 2 for a total of approximately 320 hours distributed over 8 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

A total of 24 calendar months and 6.00 person-years of professional contract labor would be required. The calendar time and level of effort required for each task and the total study are as follows:

	Schedule (Contract Months)		
Task	Task start	Task finish	Number of person-years
1. Selection of tasks and development of measures.	1	9	3.00
2. Conduct of training and collection of data.	10	17	1.50
3. Data analysis and report preparation.	16	24	1.50
	Total person-years		6.00

4. TRAINING PROGRAM REQUIREMENTS FOR ATD INSTRUCTORS

PROBLEM

The most important variables in any training situation are the ability and motivation of the students. Given just a minimum of guidance and an understanding of performance requirements, qualified students can usually progress with practice. For this reason, studies of instructor variables and differing training practices have often failed to find significant effects.

Nevertheless, it is obvious that instructors and training practices make a difference, especially in efficiency. It was apparent during the STRES effort that instructors, by and large, failed to exploit either practices or ATD capabilities in certain ways that would have promoted training efficiency. They were not aware of the complexity of discrimination and generalization processes that underlie aircrew skills nor of how feedback and guidance can be used, or misused, in teaching these processes. Potentials of ATDs in this regard were often missed because of the assumption that only an inflight training model should be used in device training. Thus the only credential generally felt to be required to teach in ATDs was aircraft certification in the skills to be taught.

As a result, ATD utilization practices were determined primarily by fidelity criteria applied task by task. If a student could not do the same things in the same way in a device that he would in an aircraft, and experience the same cues and effects, the device was usually considered of little training value. With this focus, most instructional support features would be perceived to have no significant role if their use reduced realism.

Even so, a number of instructors had become aware, usually during instructor training programs, of the value of well implemented feedback and guidance. But, they often expressed bewilderment regarding conflicting "principles" for employing these factors. And in no case was an understanding of the role of mediational processes in training apparent. If mediation were understood, device fidelity would be less an overriding concern, and instructional efficiency could be improved in all ATDs.

While the most serious problems related to ATD instructors appeared to be their lack of knowledge of learning processes and how to control them, it was also found that instructors often were unaware of many aspects of ATD use, and of the training program and how their roles related to it. At some centers, ATD instructors had undergone no formal preparation for teaching at all.

The purpose of the research outlined below is to develop an exemplar training program for ATD instructors and to identify ATD instructor

training requirements. In addition to providing an understanding of the ATD training program and their roles in it, the instructor training would enable ATD instructors better to control the learning process. They would learn to (1) recognize contingencies for using feedback and guidance and for the pacing of practice; (2) exploit mediational capabilities of students; and (3) use ATD capabilities fully to maximize training effectiveness and efficiency. In doing so, instructors would also acquire ancillary skills needed to understand task performance in terms of underlying cue-response discriminations. As a result, instructor management of learning would facilitate both ATD training and transfer of ATD skills to aircraft performance by students.

RESEARCH OVERVIEW

The study would consist of the five major tasks shown below.

Task 1: Development of an exemplar program for training ATD instructors, and preparation of an interim report.

Task 2: Conduct of ATD instructor training.

Task 3: Use of a test scenario to compare the performance of instructors who have received the experimental training with that of instructors who have not.

Task 4: Conduct of ATD training and collection of student performance data.

Task 5: Analysis of the data to determine differences in the performance of students of instructors who have received the experimental training from that of students whose instructors have not; and preparation of the final report.

ANALYTIC REQUIREMENTS

This research would be primarily an analytic effort. It would involve determining the training needs of instructors from interviews with them, observations of training practices, analyses of training problems, and reviews of training literature. Research on learning would also be reviewed. Existing summaries of uses of practice, feedback, and guidance would probably be adequate for identifying general principles of their use. However, research on discrimination and generalization processes would have to be integrated thoroughly, because contingencies for varied uses of feedback and guidance arise from these processes, and instructors must understand the contingencies to adapt feedback and guidance effectively. In addition, a heretofore ignored area, the formulation of principles concerning mediation in skill training, would have to be explored fully.

The product of the analytic effort would be an interim report describing a complete exemplar program encompassing all facets of ATD instructor roles. The major emphasis would be on preparing them to teach specified types of skills in ATDs with certain ranges of capabilities, but training would also include the understanding of instructor duties and the overall training program as a whole. The training in principles regarding practice, feedback, guidance, and mediation would be applicable to training any ATD instructor (or non-ATD for that matter), because the contingencies for their use could be translated to apply to any discrimination and generalization problem.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The purposes of the empirical part of the research would be to demonstrate that instructors trained as specified would reveal an understanding of the program, and that they would be able to use their understanding of learning processes in day-to-day training of students in ATDs. For the first purpose, instructor subjects could serve as their own controls using pre- and post-training testing to assess achievement; or a separate control group of otherwise comparable instructors without the special training could be used. (If the latter, experimental and control groups would be at different training sites.) A concurrent control group would be necessary for the second purpose, and again control instructors would not be at the same site(s) where experimental subjects would teach. To obtain sufficient numbers of subjects, the effort would probably extend over a number of months, with only a limited number of subjects available at any one time. Therefore, experimental and control instructors would be "paired" in the sense that subgroups of beginning instructors in one group would be pre- and post-tested at the same times as those in the other group.

Achievement of the first purpose would be demonstrated by written tests and evaluations by observers. Standardized scenarios depicting student performance on selected tasks would be presented to obtain one group of measures for the second purpose. These scenarios would either be simulated by computer or performed by specially trained personnel. A second group of measures would be obtained while actual students were being taught. The training regimen for this part would be standardized at all sites, and pre-training measures likely to be needed for covariance analyses would be obtained on all students to be taught by experimental and control instructors. All personnel involved in implementing the experiment and obtaining measures would be trained to fulfill their roles in a standardized manner.

Procedure

Random assignments to groups would be impractical, so all instructors would be pre-tested on knowledge of the program, and on teaching knowledge using the standardized scenario. Experimental instructors would then undergo training while the control instructors fulfilled their regular duties, including completing usual training if applicable. As the experimental instructors completed training, they and "paired" control instructors would again be tested using the standard scenario.

All instructors would then participate in ATD training of students typical of those initiating practice on the skills selected for use in the experiment. Observations would be made of instructor performances as required to obtain measures for assessing functional use of instructional principles. Measures of student performance would also be obtained that would reflect rate of ATD learning and level of proficiency at the end of ATD training.

SUBJECTS

All instructors serving as subjects should be typical of those first entering instructional roles. The students trained by experimental and control instructors should be typical of those beginning training on the skills being taught. The number of instructors should be sufficient to provide at least 14 inter-subject degrees of freedom for experimental and control comparisons; and if instructors serve as their own controls, at least 12 instructors should be used in the experimental group. The number of students should be adequate for each instructor to train at least 3 during the test.

DATA COLLECTION AND ANALYSIS

Data would be collected as explained above. Analyses would focus on (1) pre- to post-test improvement of experimental instructors, using control data for comparison if available; (2) experimental and control comparison on functional use of instructional principles; and (3) comparisons of rates of ATD learning and final proficiency levels of students taught by the two groups of instructors. Analyses of variance would be used, with repeated measures and covariance adjustments as needed.

FACILITIES

The study would be conducted in a selected CCT or CT unit and the training resources of this unit would be employed for the investigation.

The ATD would be used for a total of 520 hours. It would be employed in Task 1 for a total of approximately 15 hours distributed over 6 months, in Task 2 for a total of approximately 120 hours distributed over 8 months, in Task 3 for a total of approximately 25 hours distributed over 8 months, and in Task 4 for a total of approximately 360 hours distributed over 10 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

The research would take 30 calendar months and 6.00 person-years of professional contract labor. The calendar time and level of effort required for each task and for the total study are as follows:

Task	Schedule (Contract Months) Task start	Task finish	Number of person-years
1. Development of instructor training, and interim report preparation.	1	12	2.00
2. Conduct of instructor training.	13	20	0.75
3. Evaluation of instructors using a test scenario.	14	21	0.75
4. Evaluation of instructors using measures of student performance.	15	24	0.75
5. Data analysis and final report preparation.	22	30	1.75
	Total person-years		6.00

5. TECHNIQUES FOR EVALUATING THE PROFICIENCY OF ATD INSTRUCTORS

PROBLEM

An adequate means of evaluating the proficiency of ATD instructors is required to provide a basis for (1) diagnosis of student training problems due to ineffective instruction; (2) instructor job ratings based on teaching performance; (3) instructor incentive systems; (4) proficiency advancement of instructors undergoing instructor training; and (5) correcting or controlling for instructor effects in research on ATD learning.

At present, any data used for these purposes are collected informally and unsystematically, if at all. Some criteria used to evaluate ATD instructors are even counterproductive. For example, in some instances when teaching effectiveness has been judged by the rate of progress of students through ATD programs, some instructors have apparently advanced students with inadequate regard to their actual progress in acquiring skills. When job evaluations have been based largely or entirely on performance of collateral, nonteaching duties, some ATD instructors have, as an apparent consequence, failed to devote the effort to ATD instruction needed for adequate training.

Evaluating instructional proficiency is rife with pitfalls, and no generally satisfactory scheme for doing so has been developed, either in or outside the military. One has only to look at public education for examples of this difficulty. Nevertheless, it is a problem that should be faced. Even partial solutions would be better than no solutions at all.

The problem is that there are many ways to teach effectively and efficiently, and there are many kinds of persons who can be good instructors if relatively free to adapt teaching practices to the roles they choose to play. One requirement for any teacher using any method, however, is that he be aware of what he is doing, why he is doing it, and how well it works. He must also be motivated to succeed as a teacher, willing to assess the effects of his actions on students and adapt his methods accordingly. Finally, students must profit from the instructor's efforts.

The purpose of the research outlined below is to develop and validate techniques for evaluating ATD instructors based on the requirements just delineated. Some basic measures would be subjective, some objective. In either case, measures must be reliable, valid, and practical to use, and safeguards against misinterpretation must be provided.

RESEARCH OVERVIEW

The study would consist of the three major tasks shown below.

Task 1: Development of a candidate set of measures of instructor proficiency.

Task 2: Conduct of ATD training and use of the candidate measures to collect data concerning instructor performance.

Task 3: Analysis of the data to determine the validity and utility of the measures, and preparation of the final report.

ANALYTIC REQUIREMENTS

Techniques of measurement must derive from the manifestations of the qualities to be measured. They must also adapt to the constraints of the conditions under which observations are made. The analytic effort would be to develop the techniques according to these two criteria.

The factors of concern in teaching qualities of concern include an obvious awareness of all aspects of the training program; the amount, timing, and nature of feedback and guidance provided during ATD training; and recognition of the contingencies for their use. Recognition of the contingent value of these influences on learning, and actions taken accordingly, reveal to a great extent the instructor's knowledge of what he is doing and why. Subsequent progress, or lack thereof, by the student reveals the effects of his actions and provides contingencies for further intervention. Follow-up actions by the instructor would be indicative of his sensitivity to these effects, their meanings, and needed courses of actions.

Students would be a source for two kinds of data. Objective assessments of their progress in ATD training would be the most important type of data. In addition, their analyses of the ATD training received, anchored to problems encountered and effectiveness of training methods for overcoming them, could be very useful data.

The analytic effort would examine these possibilities as a minimum, using whatever resources that were likely to suggest measuring techniques. The product of the analyses would be a proposed measuring procedure to be validated in the empirical part of the study. The nature of criterion data would also be specified, as would the procedures for deriving evaluations from the measures.

EXPERIMENTAL METHOD

Considerations for Experimental Control

As in any validation effort, criterion variance is necessary if the discriminative capacity of measures is to be established. In the present case, optimum ranges (for validation purposes only) of instructor proficiency probably would not be available. The military justifiably eschews obviously poor instructors. Nevertheless, it was observed during the STRES study that ATD instructor proficiency does vary appreciably. In the validation of the evaluation procedures, the available range of proficiency should be represented as fully as possible.

Instructors used for the validation study would likely be teaching in a variety of programs, using a variety of ATDs in varying states of good and ill repair. Also, the number of instructors needed to validate measures may require that data be collected at given locations over a period of time so that more than one group of instructors can be used. Therefore, students who provide data should be subclassified for data purposes according to time, program, skills taught, devices used, and target aircraft, as well as to their levels of experience and/or entering competence.

All personnel involved in implementing the validation study and obtaining measures would be trained to fulfill their roles in a standardized manner.

Procedure

Prospective and criterion measures would be obtained as specified by the analytic effort. While procedures for obtaining these measures would be standardized, the instructional conditions and practices would generally be those normally characterizing instruction in the various settings.

SUBJECTS

At least 50 ATD instructors should eventually be used in the validation. They should represent a wide range of instructor competence and experience, and should be teaching in a number of different programs. At least 2 students per instructor should be used, randomly drawn from the students currently undergoing training by the instructor.

DATA COLLECTION AND ANALYSIS

Data would be collected as specified. Depending on the number of instructors involved at various times, tentative analyses would be

completed for subgroups, and/or for all data once they have been collected. The analyses would have four main thrusts. First, interrelations of instructor performance measures obtained during instruction (process measures) and those obtained from students (subjects' process and objective product measures) would be determined. Second, both product and process measures would be related to any additional criterion measures obtained. Third, because various student measures represent ... indicators of and criteria for instructor proficiency, multivariate analyses would determine within all data as a group the structural relations that reveal degree of proficiency. Fourth, from these analyses, especially the third one, patterns of data that could be used for evaluative judgments would be identified. This last step would seek patterns ... used to the extent possible of measures that could be obtained unobtrusively, reliably, and without contamination by deliberate role playing of instructors when being tested.

FACILITIES

The various training units selected for the study would supply the ATDs and other required resources. The ATDs would be used for a combined total of approximately 1,520 hours. They would be employed in Task 1 for a total of approximately 20 hours distributed over 6 months, and in Task 2 for a total of approximately 1,500 hours distributed over 8 months. It is important to note that the actual amount of ATD use in Task 2 will depend on the ATD training syllabi in the units selected. Additionally, the study would not add substantially to the use of ATDs in these units since they would be employed for routine training during Task 2.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

The research would require 18 calendar months and approximately 4.00 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

Task	Schedule (Contract Months): Task start	Schedule (Contract Months): Task finish	Number of person-years
1. Development of measures.	1	8	1.50
2. Conduct of training and collection of data.	8	15	1.50
3. Data analysis and report preparation.	14	18	1.00
	Total person-years		4.00

6. TECHNIQUES FOR THE USE OF SELF-INSTRUCTION IN ATDs

PROBLEM

The use of ATDs in continuation training (CT) is expected to increase considerably. Unlike the bulk of students now undergoing ATD training, CT aircrews will have had considerable experience in aircraft. To a great extent, they will be able to discriminate task-intrinsic feedback and to recognize the kinds of guidance needed to foster skill development and improvement. They will be able to analyze learning problems, bringing mediational tools to bear that were acquired through experience. Thus, they will be able to profit from self-instruction in ATDs in ways not possible for less experienced aircrews. Considering the volume of ATDs being procured to support CT, the effectiveness of their use in self-instruction is a matter of some practical concern.

For self-instruction to be optimal, it is necessary to design training programs that capitalize on these abilities of CT aircrew personnel, and to provide ATD capabilities necessary for skill development. The training programs must anticipate where the aircrewman needs augmented and supplemental feedback and additional guidance. They must also provide for differences in prior achievement levels of aircrews, and for the varying types of skills needed for performance of different tasks. These requirements must be fulfilled, either through automated ATD capabilities or through other instructional media.

Necessary ATD capabilities are of two general types. First, fidelity must be sufficient for realistic task performance and feedback, particularly with reference to psychomotor skills. In some instances, the experience of the aircrewman will permit substitution of mediation for physical fidelity for such skills. But, regardless of such mediational abilities of aircrew personnel, much of CT training involves the psychomotor skills and precise control coordinations required in high performance aircraft, and thus the ATDs will require considerable dynamic fidelity. A second capability that will be advantageous is the provision of guidance via automated demonstrations, and augmented and supplemental feedback. Such capability will surely be needed in some form. For example, some pilots with several years of experience reported during STRES interviews that even automated demonstrations of maneuvers designed for UPT training were valuable as standards for performance during skills maintenance training. Also, hardcopy printouts of performance records helped in analyzing sources of difficulties.

As training programs and ATDs are designed for self-instruction in a broad range of CT skills, a number of such automated instructional support features will likely be necessary. If self-instruction is to be effective and efficient, a programmatic research effort will be necessary to determine training and ATD requirements.

The purpose of the research outlined below is to develop a prototypical methodology for a programmatic effort by demonstrating (1) how the problem should be approached, and (2) how the merits of self-instruction should be validated and shortcomings identified. A comprehensive initial study of some aspects of the problem would reduce the difficulty and amount of effort required in subsequent studies.

RESEARCH OVERVIEW

The study would consist of the three major tasks shown below.

Task 1: Selection of tasks for study and development of (a) an exemplar self-instructed training program, and (b) development of measures needed to validate the program.

Task 2: Conduct of self-instructed and instructor-directed training for the same set of tasks and collection of data concerning ATD training effectiveness and efficiency.

Task 3: Analysis of the data and preparation of the final report.

ANALYTIC REQUIREMENTS

The analytic efforts would focus on task characteristics, student characteristics, and required ATD capabilities. Task characteristics would be analyzed in terms of underlying requirements for cue and response discriminations, and forms of guidance and intrinsic, augmented, and supplemental feedback needed to perform the task, to maintain it, and to teach it. Student characteristics would be analyzed in terms of variations in needs for guidance and feedback, with a strong emphasis on their mediational capabilities as related to levels and types of experiences. Also, personal characteristics such as degree of field dependency would be considered. ATD capabilities would be analyzed in terms of requirements as defined by the analyses of task and student characteristics.

The product of the analytic efforts would be an exemplar self-instruction program for training a set of tasks in ATDs with specific alternative capabilities. Types of students to be served would also be specified. Preferably, the validation of the program would involve students at different levels of proficiency so that insights might be gained regarding the value of the program for other than CT. Measures needed to validate the program would also be identified, and procedures for obtaining them specified.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The purpose of the empirical part of the research would be to validate the exemplar program and identify problems in self-instruction. Depending on the skills involved, the validation would be based on (1) demonstrated improvement in skill performance under self-instruction; or (2) achievement under self-instruction that was at least equivalent to achievement involving instructor-controlled training. In either validation case, the self-instructional program would be the same and standardized. The alternative instructor-directed program would be formalized (if necessary) to ensure all students trained in it underwent a regimen not atypical of what usually occurs. (A historical control group could be used in place of the concurrent group, but comparisons would be restricted to whatever achievement data happened to be available for them.) The range of aircrew characteristics typical of personnel undergoing the training would be represented in all groups.

All personnel involved in implementing the validation study and obtaining measures would be trained to fulfill their roles in a standardized manner.

Procedure

If a concurrent control group is used, subjects would be assigned randomly to it and to the self-instructional group. The self-instructed group would undergo training as specified, and proficiency and other measures would be obtained as planned. The control group would be trained as explained above, with similar measures obtained for concurrent control subjects. All subjects would be asked via questionnaires and interviews to critique their instruction.

SUBJECTS

If the self-instructed group are to serve as their own controls, at least 15 subjects would be used for a single homogeneous group. If student characteristics are variable enough to justify subgroups, numbers of subjects should be sufficient to provide at least 18 inter-subject degrees of freedom for any subgroup comparison. If a control group is used, at least 18 inter-subject degrees of freedom should be available for any control-experimental group comparison.

DATA COLLECTION AND ANALYSIS

Measures would be obtained as stated above, with the addition of pretraining baseline proficiency measures for self-instructed subjects if a comparable control group is not available. Analyses would focus, as

appropriate, on improvements beyond baseline performance, comparisons of rate of ATD learning between groups, and final ATD proficiency level achieved. Analyses for separate skills would reveal where self-instruction was and was not effective. Data from students' critiques would be collated and related to specific features of the alternative models of instruction in ways that revealed the appropriateness and validity of the conclusions of the analytic effort.

FACILITIES

The study would be conducted in a CT unit and the training resources of the unit would be employed in the research. The ATD would be used for a total of approximately 105 hours. It would be employed in Task 1 for a total of approximately 15 hours distributed over 6 months, and in Task 2 for a total of approximately 90 hours distributed over 6 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

Approximately 18 calendar months and 3.00 person-years of professional contract labor would be required for the study. The calendar time and level of effort required for each task and the total study are as follows:

Task	Schedule (Contract Months) Task start	Task finish	Number of person-years
1. Selection of tasks and development of training.	1	9	1.50
2. Conduct of training and collection of data.	10	15	0.75
3. Data analysis and report preparation.	14	18	0.75
	Total person-years		3.00

7. TRAINING AND EVALUATION OF ADVANCED COGNITIVE SKILLS

PROBLEM

Many of the instructors interviewed during STRES felt that advanced cognitive skills, often referred to as "headwork," "air sense," or "tactical situation awareness," were the most difficult skills to teach and that deficiencies in these skills underlay the most common and critical problems of aircrew performance. ATDs were usually not employed in attempts to train these skills. In some instances, formal training for advanced cognitive skills was simply not provided at all, even in the aircraft. On-the-job experiences in the aircraft, impromptu discussions, and other informal experiences were considered to be the main sources of training for these skills.

In recognition of this problem, research projects have recently been initiated by Air Force, Navy, NASA, and FAA research agencies to address issues associated with the training and evaluation of advanced cognitive skills. While a large number of specific research questions can be formulated concerning this problem area, some major issues are common to all problems of teaching cognitive skills. These include:

1. Development of a taxonomy for advanced cognitive skills and of methods for describing them as objectives for training;

2. Development of methods for measuring and evaluating these skills;

3. Determination of moderating factors that impact on the training of these skills such as task differences, student characteristics, and the effects of different types of stress;

4. Determination of effective and efficient methods of employing ATDs for training these skills; and

5. Development of paradigms for evaluating the effectiveness and efficiency of instruction in the skills.

The many facets of teaching aircrew cognitive skills in general, and individual ones in particular, will require programmatic efforts to ensure training effectiveness and efficiency. Nevertheless, a single intensive study to determine the utility of ATDs for teaching these skills would have implications for most if not all such training problems. To ensure generality of the results of the study, a skill involving varying degrees of complexity and cognitive requirements should be targeted so that ATD utilization could be planned, and evaluated, for a variety of skill requirements.

The research outlined below would develop and evaluate an ATD training program for the E-3A Weapons Directors (WD) crew position that focuses on teaching advanced cognitive skills associated with the task of directing air intercepts. WD training at a CT level was selected for this research because (1) the WDs perform a variety of tasks, many with extremely high cognitive loadings; and (2) the proposed research would parallel current HRL research directed toward the analysis of air intercept training for pilots.

RESEARCH OVERVIEW

The study would consist of the three major tasks shown below.

Task 1: Selection of tasks for study and development of (a) supplemental ATD training focused on the advanced cognitive skills, and (b) measures of trainee proficiency on the selected tasks.

Task 2: Conduct of ATD and flight training and collection of data describing trainee proficiency.

Task 3: Analysis of the data and preparation of the final report.

ANALYTIC REQUIREMENTS

A set of WD tasks would be selected for training that included a range of cognitive complexity and of types of cognitive skills required. ISD analyses would be performed as applicable followed by a learning analysis focusing on learning to perform under stressful conditions when contingencies for actions are constantly changing. The learning analysis would give special attention to the cue pattern discriminations required, the types of interferences among cues and responses that occur, and techniques for teaching the discrimination of cues and actions that would minimize their vulnerability to interference, especially that arising from stress itself.

The learning analysis would also identify alternative training procedures and conditions for various skills, and the skills amenable to ATD instruction would be targeted for an ATD training regimen. The regimen would include part-task segments as needed, but would emphasize realistic scenarios to be used both for training and evaluation of performance. While this regimen is under development, parallel efforts would be directed to devising procedures and instruments for assessing proficiency of WD performance. One purpose is to measure cognitive functioning in the complexity of requirements and pressures of operational performance, so the procedures must be usable in realistic settings as well as in more restricted training contexts. The measures would cover all significant aspects of the skills to be taught in the ATD, and in

addition would include where practical related skills omitted from the ATD regimen.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The purpose of the empirical part of the research would be to demonstrate the value of the ATD training for developing the cognitive skills involved, and to pinpoint any shortcomings needing correction. A control group would not be necessary provided there were other bases for inferring that desired learning did in fact occur in the ATD, and that it transfers to operational performance. (In view of the fact that formal training programs for many cognitive skills do not exist, considerable attention should be directed to identifying rate of learning variables and conditions for testing that permit experimental subjects to be used as their own "controls.")

To enhance possibilities of generalization of the analytic procedures and findings, and the results of the validation, subjects used in the empirical test should represent the full range of WDs that need further training and/or experience to gain full proficiency.

Procedure

Subjects would undergo pretesting, if needed, to establish baseline performance, including aircraft performance of related tasks omitted from the ATD syllabus. They would then receive the prescribed training in the ATD. Performance would be measured as needed to reveal rate of ATD learning and level of proficiency at the end of ATD training. To the extent feasible, all ATD trained tasks and related tasks omitted from the ATD syllabus would then be performed and assessed in the aircraft.

SUBJECTS

Subjects would be drawn from a pool representative of all WDs who would be expected either to undergo formal training for the tasks involved, or whose needed improvement in task performance would normally occur through operational experience. At least 15 subjects should be used, with the number increased as needed for multiple groups of prior experience/competence levels so as to provide at least 18 inter-subject degrees of freedom for any group comparisons.

DATA COLLECTION AND ANALYSIS

Measures would be obtained as stated above. Analyses would focus on rate of learning and level of proficiency attained by the various groups

in the ATD, using progress beyond baseline measures as dependent variables. Measures of aircraft proficiency following ATD training would similarly be compared to baseline measures in the aircraft for all groups, and on related tasks omitted from the ATD syllabus as well as on tasks included in it. If multiple groups are used, a repeated measures analysis of variance would be used to determine the differential efficacy of the ATD training for various experience/competence levels.

FACILITIES

The E-3A CCT would provide the ATD and other resources required for the study. The ATD would be employed for a total of approximately 200 hours in this study. It would be employed in Task 1 for a total of approximately 20 hours distributed over 4 months, and in Task 2 for a total of 180 hours distributed over 6 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

The study would require approximately 16 calendar months and 3.50 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

Task	Schedule (Contract Months) Task start	Task finish	Number of person-years
1. Selection of tasks and development of training and measures.	1	6	2.00
2. Conduct of training and collection of data.	7	12	0.75
3. Data analysis and report preparation.	11	16	0.75
	Total person-years		3.50

8. USE OF ATDs FOR TACTICS DEVELOPMENT AND DISSEMINATION

PROBLEM

A major frontier for the use of simulation is to support combat skills training at the CT level. In these operational units there is a unique merger between training and the development of operational tasks that are to be trained. As experienced pilots practice a combat task, they also frequently experiment with new tactics to be employed in the performance of the task. Thus, new tactics are often developed as part of the process through which experienced pilots continue to improve their skills.

In view of the capabilities of some modern WSTs, tactics development can be pursued systematically in an environment that permits safe exploration of maneuvers as well as opportunities for analyzing their strengths and weaknesses. For example, visual displays can be designed to provide geometric perspective for relations among opposing aircraft, or attack aircraft and ground targets, that are often impossible to follow in imagery. Paths of aircraft and missiles can be simulated through stochastic functions, removing the uncertainties about what an aircraft or a missile "might" be able to do. The use of airborne avionics in ATDs permits realistic practice with all the offensive and defensive aids available to a pilot in the aircraft. Alternative pathways and procedures for approaching targets can be investigated, as can various target handoff procedures. In short, many aspects of air tactics from basic maneuvering to electronic warfare can be investigated and trained in modern ATDs.

ATDs may vary in their utility as a function of whether they are used for developing, disseminating, or training tactics. For example, a laboratory device might be modifiable in ways that would not be possible with an operational ATD at a unit. Unless such a laboratory device, assuming it to be useful for tactics development, could be made available for dissemination and training purposes, little might be gained operationally. Thus, a complete system view of tactics training is required for maximum ATD utility.

Of perhaps more important concern are the research questions related to device utility as a function of classes of tactics and training requirements. For example, use of ATDs in the development of ACM tactics would require a high degree of flight dynamics and visual field fidelity, as would their use in the dissemination and teaching of such tactics. In contrast, ATD tactics development and dissemination in EW activities might make quite different fidelity demands. Thus, attention must be given to the relationships between tactical task taxonomic considerations on the one hand, and device characteristics and fidelity on the other. It is obvious, also, that the extent of simulation of the threat

environment (e.g., ground radar, AAA, SAMs, jamming, etc.) has implications for that which is feasible with ATD tactics applications.

Regardless of the classes of tactics being developed or the nature of the ATD used in their development, it will be necessary to demonstrate that such tactics developed in the ATD can be employed in aircraft. Then it will be necessary to establish the value and limits of ATD dissemination of such tactics and of their training in an ATD. These limits will probably vary, as the ATDs used for dissemination may well differ in capabilities from those of devices employed during tactics development. Also, the content and manner of ATD training for pilots in general likely should vary depending on the prior achievement levels and experiences of the pilots to be trained.

The purpose of the research outlined below is to investigate (1) the development of tactics in ATDs; (2) how tactics developed in ATDs can be validated for aircraft use; and (3) how dissemination programs can be evaluated.

RESEARCH OVERVIEW

The study would consist of the six major tasks shown below.

Task 1: Identification of general techniques for the use of ATDs in the development and dissemination of tactics.

Task 2: Development, using an ATD, of a set of tactics for a selected tactical problem.

Task 3: Validation of the tactics in an airborne simulation of the problem.

Task 4: Development of alternative methods for disseminating information describing the tactic, and use of the methods to introduce the tactic to different groups of pilots.

Task 5: Evaluation of the airborne performance of the different groups of pilots.

Task 6: Preparation of a final report describing (a) techniques for use of ATDs in tactics development and dissemination, and (b) results of the experiment.

ANALYTIC REQUIREMENTS

In developing new tactics, it will be necessary to select one or more representatives from different classes of tactics. In view of the

importance of upcoming visual system procurements, selection of a tactic requiring visual cues would seem appropriate. Ideally, the tactical area selected would be one in which there is a known need for improved performance or means of countering a threat. The details of analysis of the tactical needs cannot be specified here, but the analytic effort is critical to the developments to be investigated in the ATD. Likely, a good deal of try and revise effort would be required.

The development of the tactics would also include specifications for evaluating performance and for the contexts in which the tactics would be employed. Therefore, no additional analytic efforts would be needed for aircraft validation other than those necessary to instrument the study, obtain measures, and establish safety requirements.

The program for dissemination would require substantial analytic preparation, however. While the cue and response discriminations involved in executing the maneuvers would be identified to a great extent during development of the tactics, optimum ways to teach them ordinarily would not. Therefore, learning analyses would be completed as necessary, focusing on cues, cue interpretations, cue processing, and responses, and ways to capitalize on pilots' past learning so as to enhance acquiring the new discriminations while minimizing interference from past habits that are not conducive to performing the new skills. Problems in training resulting from differing competencies and ATD capabilities would be anticipated, and alternative teaching strategies would be developed so as to assure a viable program. Criterion measures for evaluating the performance after training would also be identified and procedures for obtaining them defined.

Of particular interest would be examination of ATD capabilities for automated demonstration/training with reference to dissemination (in the informational sense) and training. It is possible that some tactics do not require training or practice as much as they require illustration or demonstration. The ATD might fulfill such needs with auto-demo capabilities, as well as the more usual training needs.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The purpose of the validation part of the study would be to demonstrate that the tactics could be executed in the aircraft, and that they would be effective for the conceived purpose. Therefore, to the extent practical, there would be provisions for realistic targets, threats, and other aspects of the combat environment.

For the dissemination study, pilot trainees would usually serve as their own controls, i.e., baseline measures of performance would be obtained at least in the ATD for comparison with proficiency achieved

during ATD training. When practical, aircraft baseline measures would also be obtained, and/or a control group would be used who received all their training in the aircraft.

All personnel involved in instrumenting either study and obtaining measures would be trained to fulfill their roles in a standardized manner.

Procedure

For the validation, expert combat pilots would be trained to use the tactics, usually in the ATD in which the tactics were developed. Then they would attempt the tactics in aircraft, with other expert pilots filling roles of hostile forces if appropriate. Measures would be obtained as specified in the evaluation plan. In addition, the pilots trained in the ATD as well as the "hostile" pilots would be queried through questionnaires and interviews regarding their analyses of the tactics, the value and shortcomings of ATD training to perform them, and their operational applicability.

The dissemination study would begin after the tactics had been validated and the value of the ATD for training expert pilots to employ the tactics had been demonstrated. Pilots who would be expected to undergo such training, representing the full range of levels of experience and competence, would serve as subjects. They would be trained according to the exemplar program, with baseline measures and/or a control group included as explained above. Following ATD training, all subjects would perform the maneuvers in the aircraft. Measures would be obtained so as to reflect rate of learning in the ATD, proficiency level at the end of ATD training, and, as applicable, original proficiency in the aircraft upon transition and rate of progress to proficiency in the aircraft. These subjects would also be asked to critique the tactics and their training.

SUBJECTS

Only expert combat pilots, at least 3 in number, would be trained in the validation part of the study. For the dissemination study, at least 15 pilots should be used in the ATD training group if they serve as their own controls. If a control group is used, and/or if any group is subdivided by experience or competence levels, the total number of subjects should be sufficient to provide at least 16 inter-subject degrees of freedom for any group comparison.

DATA COLLECTION AND ANALYSIS

Analyses of the validation data would focus on the extent to which the tactics could be performed as designed in the aircraft.

Questionnaire and interview data would be analyzed for suggestions for improvement of the tactics and the ATD training.

Analyses of the dissemination data would focus on the transfer of the ATD skills to the aircraft, level of ATD proficiency needed to effect transfer, and the relation of ATD capabilities to skill acquisition and transfer. Questionnaire and interview data would be analyzed to identify needed improvements in the training program and equipment.

FACILITIES

An ATD with a wide angle visual system would be employed in the study. The ATD should be configured specifically as the aircraft in which the tactic is validated. An ACMI would also be employed in the study. The ATD that is selected for this study would be employed for a total of approximately 130 hours. It would be employed in Task 1 for a total of approximately 20 hours distributed over 2 months, in Task 2 for a total of approximately 50 hours distributed over 3 months, and in Task 4 for a total of approximately 60 hours distributed over 6 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

The study would take 18 calendar months and 4.00 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

Task	Schedule (Contract Months) Task start	Task finish	Number of person-years
1. Identification of techniques for tactics development and dissemination.	1	4	0.75
2. Development of tactics.	4	6	0.75
3. Validation of the tactics.	7	8	0.50
4. Development and use of methods to disseminate the tactics.	9	14	0.75
5. Evaluation of pilot performance.	14	17	0.50
6. Preparation of the report.	16	18	0.75
	Total person-years		4.00

9. TECHNIQUES FOR CREW INSTRUCTION

PROBLEM

Crew training in aircraft is expensive, and often difficult to conduct because of constraints imposed during flight. Although ATDs offer unique training capabilities that can be employed to improve the efficiency and effectiveness of crew training, these capabilities are frequently not exploited due to the use of the inappropriate inflight training model during crew instruction in ATDs.

While there has been some research on team processes and team training, relatively little is known concerning how best to utilize ATDs for such training. For ATDs to be used effectively in crew training, there must be systematic investigation of the dimensions of crew behavioral processes and the learning problems related to such processes, and of the means to manage, manipulate, and control the learning of such behaviors through ATD use.

In order to achieve effective ATD usage in this area, a number of questions need answering. For example, there is need to determine which tasks require crew training in addition to individual position training. Further, there is the question of how best to integrate crew and individual training in the ATD and aircraft. There is the rather fundamental question of whether principal crew processes (e.g., coordination, compensation, communication) are really trainable in the sense of deliberate management of the instructional process; there is the possibility that such crew process skills just "grow" or "happen" as a by-product of practice and experience and are not directly manageable through training intervention. However, assuming optimistically that they are trainable in the process management sense, training objectives must be formulated and ATD training experiences designed to produce the desired learning. Finally, assuming trainability, there are the specific instructional questions such as how to provide feedback and guidance to the team in the ATD, how best to use cognitive training, how to sequence and pace training, and the like.

The purpose of the research outlined here is to develop an exemplar program for training one type of aircrew in an ATD and to validate the program empirically. The validation would focus on the feasibility of implementing the program operationally, and on training effectiveness as exhibited in crew performance in aircraft. To ensure applicability of general instructional techniques in developing programs for other types of crews, the exemplar crew tasks would be chosen so as to represent a variety of crew interactions.

RESEARCH OVERVIEW

The study would consist of the three major tasks shown below.

Task 1: Selection of crew tasks to be studied and development of (a) an exemplary crew training program, and (b) measures of crew performance; preparation of interim report.

Task 2: Conduct of ATD and flight training and collection of performance data for crews trained in the experimental programs and in an existing program.

Task 3: Analysis of the data and preparation of the final report.

ANALYTIC REQUIREMENTS

The intent is to develop a program that would maximize the use of ATDs in crew training. An aircraft system would be selected that would present the required variety of crew interactions. Factors to be considered in the selection would include crew complexity (e.g., F-4 vs. C-5) and accessibility of crew to observation and data gathering. ISD analyses would be completed for the selected system, as applicable, to identify team tasks. This would be followed by a learning analysis to identify essential discriminations in the team tasks, those amenable to ATD instruction, and effective methods and ATD requirements for teaching them and ensuring generalization to aircraft performance. Existing techniques for crew training would be surveyed, focusing not only on aircrew training, but team training in other military contexts and in industry. Measures of crew performance would also be developed that revealed separately the proficiency of individual crew members and the proficiency of the aircrew as a unit. (An outline for the development and validation of such a measure is included in this report. It would not be necessary to wait for that project to be completed, however, for existing, less thorough techniques of assessing individual and crew proficiency could be adapted to the needs of this study.)

The product of the analytic effort would be an interim report specifying skills to be taught, underlying discriminations to be achieved, procedures for teaching them using a designated ATD, and techniques for assessing training effectiveness.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The validation of the ATD aircrew training program would consist of an empirical investigation of its effectiveness and efficiency. Therefore, all factors not relevant to the implementation of the program,

but which could influence the quality of aircrew performance, should be controlled either through manipulation of conditions and random assignment of crews to concurrent experimental and control groups, or through covariance adjustments based on crew characteristics that differ from crew to crew. All personnel involved in implementing the validation study and obtaining measures would be trained to fulfill their roles in a standardized manner.

Procedure

All crews would be assigned randomly, if practical, to concurrent experimental and control groups. If a historical control group is used, or if it is not feasible to assign crews randomly, measures of crew characteristics would be obtained as needed for covariance analyses. The control group would be trained in an aircraft and associated ATD according to an existing regimen. The experimental group would receive training in an ATD according to the newly devised program. The experimental group would transition to the aircraft when and as specified in the exemplar program. Measures would be obtained for both groups to reflect rate of learning in the ATD, level of proficiency at the end of ATD training, original level of proficiency in the aircraft, and rate of progress to full proficiency in the aircraft.

SUBJECTS

Crews should be typical of those who would initiate crew training of the sort covered by the exemplar program. At least 10 crews should be assigned to each group. If the experience levels of crews or individual crew members are heterogeneous, experimental and control groups should be subdivided into homogeneous levels, with enough subjects per group to provide at least 18 degrees of freedom for any inter-group comparisons.

DATA COLLECTION AND ANALYSIS

Measures would be obtained as stated above. Multiple group, repeated measures analysis of variance would be used, with different skill measures and times they were obtained comprising two correlated dimensions. Covariance adjustments would be included as needed for between-group comparisons.

FACILITIES

The ATD and other resources of the CCT unit selected for the study would be employed for the research. The ATD would be used for a total of approximately 230 hours in this study. It would be employed in Task 1 for a total of approximately 30 hours distributed over 4 months, and in Task 2 for a total of approximately 200 hours distributed over 8 months.

SCHEDULE

The study would take 24 months to complete. Approximately 4.00 person-years of professional contractor labor would be required. The calendar time and level of effort required for each task and the total study are as follows:

Task	Schedule (Contract Months) Task start	Task finish	Number of person-years
1. Selection of tasks, development of training and measures, and interim report preparation.	1	12	2.50
2. Conduct of training and collection of data.	13	20	0.75
3. Data analysis and report preparation.	18	24	0.75
	Total person-years		4.00

10. TECHNIQUES FOR EXTENDED TEAM TRAINING

PROBLEM

Aircrews and personnel in different systems must coordinate their actions during combat. This "extended" team varies in composition from a strike aircraft pilot coordinating a bomb run with a forward air controller, to pilots of multiple aircraft being coordinated by a combat center at a remote airborne or ground location. Training extended teams is expensive in the actual systems, and difficult because of complex logistics and coordination problems, lack of control over training, and concern for safety and exigencies of flight.

Extended team training is also difficult to conduct in existing ATDs. Indeed, most ATD training programs do not include extended team training other than as part of routine communication tasks. If interactions with remote members of an extended team are included in training, the instructor usually must simulate these interactions himself. The instructor is frequently burdened with other instructional tasks and may lack sufficient knowledge of the other team system to simulate it appropriately. As a result, the quality and frequency of simulation of remote team members is often degraded. Moreover, since inflight training models are usually employed, there are no attempts to exploit the potential for training effectiveness and efficiency offered through simulation.

Programmatic research is needed to determine (1) effective and efficient techniques for teaching extended team tasks in ATDs, and (2) design requirements for ATDs for extended team training.

The research outlined below focuses on the development of effective techniques for the use of ATDs in the instruction of extended teams. The analytic efforts involved would by necessity address design requirements and manners of use of ATDs. The task of training two fighter pilots in separate aircraft to work with one another in air combat engagements has been selected for prototypical development because adequate facilities now exist for the study of this type of task coordination. Also, it is a critical tactical training area for which little is known concerning how to exploit the unique training capabilities of ATDs.

RESEARCH OVERVIEW

The study would consist of the three major tasks shown below.

Task 1: Selection of extended team tasks to be studied and development of (a) an experimental extended team training program, and (b) measures of extended team performance.

Task 2: Conduct of extended team training in the ATD and aircraft and collection of performance data.

Task 3: Analysis of the data and preparation of the final report.

ANALYTIC REQUIREMENTS

ISD analyses would be performed as applicable, followed by a learning analysis to relate the interactions of the pilots to essential discrimination, coordination, and interference problems involved in team learning and performance. Reviews of literature and interviews with pilots experienced in extended team performance, and with pilots undergoing training, would help identify task requirements and current training problems and shortfalls. A training regimen would be developed to optimize team training, including the accommodations of interferences arising within each aircraft and in coordinated efforts. Measures of proficiency of the extended team performance would be developed that focused on the adequacy of coordination of actions as well as of the actions themselves.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The purpose is to demonstrate that extended team training occurring in ATDs transfers to performance in aircraft. Therefore, no formal control group is required, provided the amount of effort currently needed to train in aircraft is, or can be, known. If such information is not available, then a control group must be trained in the aircraft and their progress monitored as for the experimental group.

Procedure

Subjects would be assigned randomly to an experimental and a control group (if used). If a historical control group is used, measures of subject characteristics likely to affect training results would have to be available in the files of control subjects, and obtained at the outset for the experimental subjects. The experimental groups would be trained as specified by the newly developed program. Measures would be obtained to reflect rate of learning in the ATD, level of proficiency at the end of ATD practice, original level of proficiency in the aircraft, and rate of progress to full proficiency in the aircraft.

SUBJECTS

Subjects should be typical of those initiating training of the sort covered by the new program. It is assumed that they will have previously

mastered the individual piloting skills relevant to the aircraft of concern. If they are heterogeneous in levels or types of experience, they should be subdivided into homogeneous groups for comparisons of results. If only a single experimental group is used, at least 15 pairs of pilots should be involved, with the number increased as needed to provide at least 16 inter-pair degrees of freedom for any subgroup comparison. If a concurrent control group is used, each individual group should have at least 10 pairs, and any between-group comparison for subdivided groups should have at least 16 inter-pair degrees of freedom.

DATA COLLECTION AND ANALYSIS

Measures would be obtained as stated above. Analyses would focus on rate and level of ATD learning and original level of proficiency in the aircraft and rate of progress to full proficiency in the aircraft. The two learning rates and levels would be compared task by task for experimental and control groups, using repeated measures analysis of variance as appropriate for total and/or subdivided groups.

FACILITIES

Two full FOV visual ATDs configured as fighter aircraft would be required for the study. These ATDs must be capable of being coupled together for joint ACM training. Candidate facilities include (1) NASA Langley DMS, (2) McDonnel Douglas MACS, and (3) AFHRL/OT SAAC. The ATDs that are selected for this study would be employed for a total of approximately 230 hours (summed over both ATDs). They would be employed in Task 1 for a combined total of approximately 30 hours distributed over 4 months, and in Task 2 for a combined total of approximately 200 hours distributed over 10 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

A total of 20 calendar months would be required for the study. Approximately 4.00 person-years of professional contract labor would be required. The calendar time and level of effort required for each task and the total study are as follows:

Schedule
(Contract Months)

Task	Task start	Task finish	Number of person-years
1. Selection of tasks and development of training and measures.	1	8	2.00
2. Conduct of training and collection of data.	9	18	1.00
3. Data analysis and report preparation.	17	20	1.00
	Total person-years		4.00

11. OPERATIONAL TRAINING PROGRAM DEVELOPMENT

PROBLEM

There have been relatively few instances in the military in which comprehensive ATD training programs were developed following a training model, as opposed to a flight model, approach. Systematic procedures such as ISD have generally not been applied to ATD training programs. Even so, when they are applied, ISD and similar approaches do not exploit all learning technology has to offer to ATD training program design. While the need to learn cue discriminations, for example, is recognized and some provisions for their learning are included, these systematic approaches do not address the complexities of acquiring discriminations. There are no comprehensive provisions for anticipating interferences within the learning processes themselves, nor for distinguishing among types and patterns of discriminations of cue and response dimensions that can be generalized readily when needed to transfer skill performance from an ATD to an aircraft, or even to appreciably different situations within the same vehicle. Mediational foundations of transfer are especially slighted, and general rules for providing guidance and feedback do not address specific difficulties in acquiring generalizable discriminations that are invulnerable to interference. In fact, the routine dependence on proficiency criteria ignores possible future interferences that must be accommodated for superior retention.

There is a system-wide need to develop ATD operational training programs that consider all facets of the learning process as they relate to the development of stable, enduring skills. Equally important, skills and the underlying cue and response discriminations developed in one program should systematically anticipate, and provide for, building more advanced skills during training subsequent to the program. Current training programs, including portions developed using ISD, address this aspect of transfer only intuitively if at all.

The purpose of the research outlined below is to develop an exemplar operational training program that takes all of these factors into account. The primary effort is analytic, but an empirical demonstration of the value of the product would be made. It is through such demonstrations that the program gains credence, and the value of training media used is established. Evaluation of the effectiveness of the operational ATD training program would also provide information concerning the effectiveness of instructional techniques and ATD design features employed in that program. Furthermore, the success of an operational program also indicates, although indirectly, the effectiveness of the techniques used to develop the training and design the ATD. Thus, knowledge would be gained concerning the effectiveness of the systematic application of training technology to ATD instruction and design.

RESEARCH OVERVIEW

The study would consist of the three major tasks shown below.

Task 1: Development of specifications for the conduct of the operational training program; preparation of interim report.

Task 2: Conduct of training, and collection of training effectiveness and efficiency data, and revision of training program.

Task 3: Analysis of the data and preparation of the final report.

ANALYTIC REQUIREMENTS

A relatively comprehensive ATD training program would be selected so that a wide variety of training problems would be included, and validations of the program in actual use could address the training strategies derived for the variety of problems. Thus, adaptations of analyses could be made readily to a number of other ATD programs with related problems, and occasional shortcomings in the first effort could be corrected with predictable results.

Thorough ISD analyses would be completed as needed and applicable. Literature reviews would seek to identify the full range of cues needed for the tasks to be taught. (A number of other experiments outlined in this report, particularly those concerned with ATD instruction of operational tasks, would provide many useful insights and data in this regard.) The ISD analysis would be followed by a thorough learning analysis addressing issues such as those identified above. Also, reports describing development of training programs by commercial airlines should be consulted, for, all in all, they have made more extensive use of multi-media resources and alternative training strategies.

The final product would be an interim report with specifications for the conduct of the operational training program, including training strategies and media and schedules for their use. Additionally, requirements for all other aspects of the aircrew training program would be specified as necessary to ensure proper preparation of the trainee and instructor for ATD training, and exploitation of ATD trained skills in the aircraft. All guidance pertaining to various aspects of ATD training provided in the Utilization report of STRES would be heeded.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The purpose of the empirical part of the study would be to demonstrate the value of the exemplar program, which may or may not involve comparisons with alternative programs. Thus, the primary concerns of experimental control are (1) that the program be implemented as conceived; (2) that all levels of students to be served by the program be included; and (3) that achievement measures and other data be collected in ways and at times that permit diagnostic as well as overall evaluations. Preparation of instructor and measurement personnel would be crucial, of course, but all personnel who are involved in managing and implementing the program should be fully aware of its nature and their responsibilities.

Procedure

The program would be implemented as designed. Student progress and instructional practices would be monitored regularly. Measures and other data would be collected as needed to identify student personal and background characteristics; rates of learning, separated by skills and media/phases/stages of training; proficiency levels achieved in the ATD and aircraft; and cost benefits.

SUBJECTS

Subjects would be students normally entering the training covered by the program. Data would be obtained from successive classes as available, and as needed to confirm the value of the program and identify shortcomings to be corrected.

DATA COLLECTION AND ANALYSIS

Measures and other data would be collected as indicated above. Analyses would provide summary statistics revealing ATD training efficiency and effectiveness of transfer to the aircraft. Additional detailed analyses would focus on strengths and weaknesses related to individual skills, media, training strategies, and individual student characteristics. Also, overall cost analyses would be made. If students trained previously under an old regimen are to be used as a historical control group, comparisons of their achievements with those of students in the new program would be made for those aspects where historical data were available.

FACILITIES

The ATDs and other training resources of the unit selected for study would be employed. If it is determined that the existing ATD employed in the program selected for study is deficient in design, then a laboratory ATD, such as the ASPT, would be configured to remedy the deficiencies. The simulator that is selected for this study would be employed for a total of approximately 2,050 hours. It would be employed in Task 1, if available, for a total of approximately 50 hours distributed over 5 months, and in Task 2 for a total of approximately 2,000 hours distributed over 13 months. The actual amount of use of the simulator in Task 2 would depend on the training syllabi developed during Task 1 and the flow of students through the program.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

If an existing ATD is employed, the study would take approximately 36 calendar months to complete and require 8.00 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

Task	Schedule (Contract Months) Task start	Task finish	Number of person-years
1. Development of training and interim report preparation.	1	18	4.50
2. Conduct of training, collection of data, and revision of training program.	19	31	2.00
3. Data analysis and report preparation.	30	36	1.50
	Total person-years		8.00

12. ATD DESIGN REQUIREMENTS AND TECHNIQUES FOR TEACHING SELECTED OPERATIONAL TASKS

PROBLEM

With the advance of simulation technology, opportunities for simulator training of aircrews have expanded appreciably. The use of ATDs in CCT and CT programs especially appears likely to result in considerable savings in cost. At the same time, many of the constraints imposed in aircraft training by environmental factors and safety requirements can be reduced or removed, permitting more efficient and effective training.

If the full value of simulation is to be realized, ATDs and the CCT and CT programs in which they are employed must be designed to capitalize on the unique contributions ATDs can make to skill learning. Generally, this means that an optimum training model should drive ATD and program design. Such a model would usually be different from the aircraft model of task performance. At present, guidance for the development of training models for these expanded uses of ATDs is rudimentary at best due to the lack of experience, systematic or otherwise, with the use of the latest simulation technology in the instruction of operational (CCT or CT level) tasks.

To address this deficiency, the research proposed here would generate knowledge concerning the design and use of ATDs and provide specific recommendations for ATD design and training requirements for selected operational tasks (e.g., teaching specific ACM maneuvers to F-16 pilots). The immediate goal of the research would be to develop models for the use of ATDs that could be employed in the CCT or CT programs in which the particular tasks investigated are taught. The ultimate goal however, would be the development of general guidance for the design and use of ATDs, and the assessment of the validity of the application of this guidance.

The research would be conducted utilizing a wide variety of operational tasks to ensure a broad basis for the development of the guidance. Tasks selected for study would be (1) those currently being trained inefficiently, if at all, in existing ATD programs; or (2) those tasks that have not been previously trained in ATDs, but will soon be trainable in the newest generation of ATDs (e.g., ACM). The STRES high value tasks include examples of such tasks. The following plan describes a prototypical experiment for a selected task.

RESEARCH OVERVIEW

The study would consist of the four major tasks shown below.

AD-A094 625 CANYON RESEARCH GROUP INC WESTLAKE VILLAGE CALIF F/G 5/9
SIMULATOR TRAINING REQUIREMENTS AND EFFECTIVENESS STUDY (STRES)--ETC(U)
JAN 81 W W PROPHET, J B SHELNUTT, W D SPEARS F33615-77-C-0067
UNCLASSIFIED AFHRL-TR-80-37 NL

2 OF 2
AD-
A094625

END
DATE
FILMED
3 81
DTIC

Task 1: Development of experimental, exemplar models for ATD design and use for training the selected tasks.

Task 2: Configuration of the ATD to meet study requirements, and development and conduct of ATD instructor training.

Task 3: Conduct of experimental training for one group of trainees and of conventional training for another group; collection of trainee performance data in the ATD and aircraft.

Task 4: Analysis of the differences in ATD training effectiveness and efficiency between the two groups, and preparation of the final report.

ANALYTIC REQUIREMENTS

The value of the project would depend entirely on the thoroughness of the analytic effort in developing the program. All aspects of ISD processes would be involved as applicable. In addition, learning analyses would pick up where usual ISD efforts stop. Specifically, it would not be sufficient to identify cues to be discriminated, responses to be associated with them, and contexts for task performance, or to identify only general provisions for guidance and feedback. Learning analyses would be required to identify problems of learning the discriminations involved; patterns and variations in task presentations that would circumvent the problems; facilitative mediational competencies to be built through judicious sequences of learning tasks; contingencies for using and adapting guidance and feedback; specification of generalization structures to be taught to ensure transfer; etc.

The products of the analytic efforts would be the identification of a set of skills to be taught and an exemplar model for training them in given ATDs. Additionally, ATD design requirements for the conduct of this training would be specified. (If an ATD already exists for the task of interest, it may be beneficial to use that ATD and simply attempt to maximize effectiveness and efficiency of training of the task in the existing device.)

EXPERIMENTAL METHOD

Considerations for Experimental Control

Two possibilities for control groups, or a combination of them, should be considered. First, the skills selected for training program development may be currently taught only in aircraft. Second, training on at least some of the skills may currently be provided in ATDs. In either case, aircrews trained by current practices could serve for

control comparisons, with as many control groups (including historical) as needed to cover current practices. Training practices for these groups would not be changed, except as needed to ensure that all subjects in them received a standard regimen, and that all tasks to be learned by the experimental group(s) were included at some stage.

Instructors would be trained to conform to the specified regimen, and measurement personnel would be trained to obtain measures in a standardized manner.

Procedure

Subjects would be assigned randomly to concurrent control and experimental groups. If historical control subjects are used, preliminary measures on current subjects would be obtained comparable to those in the files for the historical groups. Each control group would receive a standardized form of current training. The experimental group would receive training specified as most appropriate in the training model. All subjects undergoing ATD training would transition to the aircraft following the schedule specified in the model. Measures and other data would be obtained that would permit assessments of rate of learning in the ATD and of progress to proficiency in the aircraft; level of achievement at the end of ATD training; original level of proficiency in the aircraft; and cost effectiveness.

SUBJECTS

Subjects would be aircrew students at the stage they should normally enter training on the skills of concern. If they are heterogeneous in experience, either they should be drawn from a homogeneous part of the subject pool, or preferably, all levels of experience would be included but controlled by subdividing experimental and control groups. The total number of concurrent subjects should be sufficient to provide at least 18 inter-subject degrees of freedom for any comparison.

DATA COLLECTION AND ANALYSIS

Measures would be obtained as appropriate for the tasks being trained, and on a schedule permitting the assessments identified above. In addition, data needed for cost estimates would be collected following standard guidelines. Training effectiveness and efficiency would be determined through inter-group comparisons using multiple group repeated measures analyses of variance. (Covariance adjustments for student characteristics would be made if historical control groups are used.) Cost analyses would normally focus on only those items that differed among groups. (It may be desirable for general management purposes, however, to conduct a comprehensive cost analysis of at least the exemplar program.)

FACILITIES

A laboratory ATD, such as the ASPT, would be configured to meet study requirements and would be employed for a total of approximately 120 hours. It would be employed in Task 1 for a total of 20 hours distributed over 4 months, in Task 2 for a total of 30 hours distributed over 2 months, and in Task 3 for a total of 70 hours distributed over 4 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

Given the existence of an appropriate ATD, the study would take 18 calendar months and require 3.00 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

Task	Schedule (Contract Months) Task start	Task finish	Number of person-years
1. Design of the ATD and development of training.	1	8	1.50
2. Configuration of the ATD and development and conduct of instructor training.	5	12	0.50
3. Conduct of training and collection of data.	13	16	0.25
4. Data analysis and report preparation.	15	18	0.75
	Total person-years		3.00

13. DEVELOPMENT AND VALIDATION OF A MODEL FOR PREDICTING ATD TRAINING EFFECTIVENESS

PROBLEM

Decisions regarding the design and utilization of ATDs should be made with a knowledge of the potentials and limitations of alternative device capabilities for effective training. Such decisions now are based largely on the assumption that effective ATD training can be achieved only to the extent devices provide stimuli and permit tasks to be performed just as they are in a target aircraft. As a result, physical fidelity of ATDs to aircraft has been the criterion for acceptable designs of the devices and for allocating training to them.

It became apparent during the STRES study just how restrictive this assumption has been. The potentials of many ATDs were not exploited because an aircraft training model could not be adapted to them with the degree of realism assumed necessary. The impact of this assumption on design requirements is revealed by the statement of one researcher who stated that his goal was to have a pilot emerge from a simulator saying that it felt exactly like the aircraft simulated.

If a valid technique were available for predicting the training effectiveness of devices, decisions regarding their utilization could be based on a realistic assessment of their training values rather than on the assumed requirement of physical fidelity. Formalized as a model, the technique would also provide predictions of training value for alternative device characteristics that would be of value when design decisions are made. In fact, a comprehensive model for predicting training effectiveness of ATDs would make possible numerous informed decisions that now are more or less guesswork. If the effectiveness of given training practices using an ATD with given capabilities could be estimated reliably, tasks could be allocated to devices for training with clear understanding of the likely outcome. Inadequacies of devices could be identified beforehand, permitting supplementary training needs to be pinpointed. In brief, if one knew in advance what an ATD feature could accomplish in training, he would know how to design the device and how to make optimum use of it.

A few attempts have been made to develop models for predicting the effectiveness of training devices. None has been of practical use, and for a very simple reason: They all focus on the physical characteristics of the device. How the devices are used in instruction is a secondary consideration if indeed it is even included as a factor. Proposed models also have ignored the mediational capabilities of students who might substitute imagery for physical and task fidelity. They have used simplistic notions of transfer of training and very superficial mathematical analyses in formulating equations for predicting training effectiveness.

The research outlined below would develop a model for predicting training effectiveness of ATDs that incorporates, and formulates in a reasonable manner, the complex processes of skill learning. It provides for validating the model using various types of ATDs. Five types of factors would be represented in the model: (1) fidelity dimensions of device characteristics and capabilities; (2) skills underlying performance of individual tasks; (3) adequacy of instructional practices; (4) mediational capabilities of students as inferred from their past experiences; and (5) a logically consistent and coherent conception of transfer of training. The model would differ substantially from those previously developed because of realistic appraisals of the last three factors.

RESEARCH OVERVIEW

The study would consist of the four major tasks shown below.

Task 1: Development of a model for predicting ATD training effectiveness; preparation of an interim report.

Task 2: Conduct of training in the ATD and aircraft and collection of data concerning ATD training effectiveness.

Task 3: Analysis of the relation between predicted transfer of ATD-trained skills to the aircraft and actual manifestations of the transfer.

Task 4: Preparation of a final report that describes the model and the results of the validation study.

ANALYTIC REQUIREMENTS

The research is primarily analytic in nature, and its success would depend entirely on the thoroughness with which the model incorporates the complexity of learning processes and device characteristics identified above. At the same time, the model should readily be adaptable to practical applications under conditions in which predictive assessments would be made. Some compromises between thoroughness and practicality would be necessary, but under no condition should the current penchant for simplicity be the criterion for making them.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The first four of the five types of factors to be represented in the predictive model are each of procedural concern in the validation of the

model. Each of these four types has countless numbers of possible variations. However, these variations pose no difficulty for validation so long as (1) two or more realistic, representative variations within each type can be included in the validation experiment; and (2) the model permits quantifications of the differences between variations within a type.

Thus, the validation experiment would be a four-factorial design with at least two <u>divergent</u> levels for each factor as indicated below:

	Factor	Levels
I.	Device/task fidelity	Relatively low vs. relatively high
II.	Consonance of task requirements to device physical characteristics	Conditions for task performance fulfilled objectively by device design vs. partial fulfillment only
III.	Instructional practices	Training designed for optimum use of feedback, guidance, and mediation, using available instructional features vs. less than optimum (but common) uses of these influences and features
IV.	Relevance of students' experiences	Could be varied directly, but more economical in this original validation to vary in relation to both types of tasks in Factor II

Note that a variety of representations of each factor would be acceptable, so long as there is variation on each, and each factor is measured. The model would provide *a priori* the procedures for measurement based on analytic considerations. Additional factors to be held constant during the validation experiment would be any influences (e.g., equipment functioning, periodicity of practice) that would introduce irrelevant effects on measures.

Procedure

Subjects would be assigned randomly to as many groups as levels (types) of ATDs used. A small number of tasks would be trained, with half at each level for Factor II above. Each task would be practiced on each device by all students using that device. Instruction appropriate for each task would be standardized within levels of Factor III, and instructors would be trained to follow the standard regimens. Rates of progress during ATD training and level of achievement just prior to transition to an aircraft would be measured. All students would transition to the aircraft immediately upon completion of ATD training, and

initial aircraft performance level and rate of progress to proficiency would be measured.

SUBJECTS

Subjects would be student pilots with comparable training and experience backgrounds. The total number of subjects would be determined by the number of levels used for Factors I and III above. Assuming only two levels of ATDs (I) and two of instructional practices (III), a total of 32 subjects, 8 per subgroup, would be needed. With additional levels, 6 subjects per subgroup would be adequate.

DATA COLLECTION AND ANALYSIS

A necessary minimum of instructors would be trained to obtain the measures in a standardized manner. If preliminary reliability checks on any measure reveal inter-instructor reliabilities below .75, the average of measures obtained by at least two instructors would be used for analyses involving those measures.

Data required to implement prediction via the model would be acquired before the validation experiment was begun. The model would specify the data needed and how it would be quantified and reduced. Periodic measures of performance in the ATD and in the aircraft would be at intervals adequate to determine rate of skill acquisition. Levels of ATD performance at the end of ATD training and in the aircraft immediately following transition would be assessed at these times

Analyses of data would focus on the relation between predicted transfer of ATD-trained skills to aircraft and actual manifestations of the transfer. At this stage, the predictions would have to be in relative terms. There would be no way a priori to calibrate measurement units in the model against units of performance that must vary with tasks and how proficiency is assessed in aircraft. Relative rankings of predicted transfer, and proportionate metric differences between these ranks, would be compared with rankings of manifest transfer and their proportionate difference metrics. These comparisons would be made task by task for homogeneous subgroups as defined by levels on Factors I and III. Within the scope of this design, tests of statistical significance would not be needed, although confidence intervals of measures of manifest transfer should be computed so that distinctness of subgroup differences could be determined.

FACILITIES

No special facilities would be required for the analytic part of the study, although access to a variety of library materials would be

necessary. For the validation experiment, two or more ATDs would be required that are alternative designs for the same training program. It may be possible to employ alternative F-16/A-10 or B-52 WST designs in the study. Additionally, laboratory ATDs configured to meet specific study requirements could also be employed. The ATDs that are selected for the study would be employed in Task 2 for a combined total of approximately 320 hours distributed over 8 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

If existing ATDs are employed in the study, the research would take 30 calendar months and require 6.00 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

Task	Schedule (Contract Months) Task start	Task finish	Number of person-years
1. Development of a model and interim report preparation.	1	12	2.50
2. Conduct of training and collection of data.	13	20	1.00
3. Data analysis	18	24	1.50
4. Final report preparation.	24	30	1.00
	Total person-years		6.00

14. DESIGN REQUIREMENTS FOR ATD USE IN PROFICIENCY EVALUATION

PROBLEM

The assessment of aircrew proficiency is an expensive and often quite difficult endeavor. Aircraft are used for those assessments as a rule, but a number of factors make it desirable to use ATDs where practical. In addition to the cost of aircraft operation, these factors include the expense of ranges and other facilities employed to provide combat environments, safety considerations that restrict what can be done, and where, in aircraft, and the need to conserve energy. Furthermore, objective and comprehensive measures of individual and system performance, particularly for certain combat skills, are often difficult, in some cases impossible, to obtain in aircraft. And for measures that can be obtained satisfactorily, reliabilities and meanings are frequently compromised because of uncontrollable variations in the weather, atmospheric conditions, and other aspects of the environment.

When simulators can be used, as they have been for certain emergency procedures evaluations, instrument checks, and some tactical checks, less expense is generally involved, and the difficulties imposed by the other factors are reduced as well. Similar benefits will accrue as the use of ATDs is extended to other types of proficiency assessments. Also, assessments of some skills which cannot generally be performed at all under realistic conditions in aircraft may well be made with more or less completely simulated realism.

Expanding the uses of ATDs for proficiency checks requires assurance that evaluations made in ATDs would be essentially the same were they made in aircraft. Specifically, if an aircrewman would pass the check for a given task in an ATD, he would also pass it in an aircraft, and if he would fail in an ATD, he would also fail in the aircraft.

It is crucial that both conditions obtain if ATDs are to be used to assess proficiency of aircraft performance. (Some studies at least tangentially addressing this problem have not provided adequately for both conditions, either because of inadequate experimental designs or inappropriate analyses of data.) Prerequisites for valid assessments in ATDs are device capabilities and characteristics that permit the fulfillment of these two conditions to be determined.

Considerable programmatic research will be necessary to relate required device capabilities and characteristics to validity of proficiency assessments on a task by-task basis. Two types of studies could be undertaken concurrently. One type would require relatively little prior preparation and adaptation of existing ATDs. It could also be coordinated with ongoing training and proficiency maintenance programs. Specifically, measures of task proficiency would be obtained for selected tasks as they are performed in an ATD, using as subjects the aircrews

normally participating in the ATD program. As soon as practical, comparable measures would then be obtained during performances of these tasks in aircraft. Intercorrelations of the various measures would reveal the extent to which aircraft proficiency could be predicted from performance in the particular ATD used. As a data base accumulated that included similar studies using different tasks and types of aircrews, patterns of intercorrelations could be examined as they related to task requirements and ATD capabilities.

The second type of study would involve formal experiments. It would have the advantage, however, of arriving at definitive answers sooner. The strategy for the experiment would be to (1) identify all cues and responses normally involved in aircraft performance of the tasks involved; (2) provide, to the extent practical in an ATD, these cues and opportunities to perform all required actions in a realistic manner; (3) determine baseline correlations among measures of ATD performance levels and those in aircraft; (4) remove or degrade cues in the ATD and its performance dynamics to determine the effects on the baseline correlations; and (5) identify minimum ATD cues and dynamics needed for adequate prediction of aircraft performance on separate tasks from performance in the ATD.

RESEARCH OVERVIEW

The study would consist of the five major tasks shown below.

Task 1: Identification of device capabilities and characteristics potentially related to the use of ATDs for proficiency evaluations and selection of feasible device factors for study.

Task 2: Selection of skills to be assessed and development of procedures for assessing them using the ATD and aircraft.

Task 3: Assessment of the skills using the ATDs and aircraft. Device factors identified in Task 1 will be varied as appropriate.

Task 4: Analysis of the data to determine the contribution of each ATD design factor to prediction of aircraft performance.

Task 5: Preparation of a final report that described (a) all device capabilities or characteristic factors analytically determined to be related to the use of ATDs for proficiency evaluation, and (b) the results of the experiment.

ANALYTIC REQUIREMENTS

Tasks to be used in the experiment would be identified as well as cue and response requirements for each task. An ATD would be selected and adapted as necessary so as to provide for these cue and response requirements in as realistic a manner as practical. Measures of task proficiency would be adapted from existing proficiency check criteria for use in the ATD and the aircraft. Additional measures would be developed for aspects of performance that are likely to correlate with proficiency either in the aircraft or the ATD. ATD measures should be more detailed and comprehensive than those used in aircraft so as to provide as broad a base of candidate predictors as practical. Also, measures would be automated to the extent feasible.

Special attention would be given the problem of reliability of criterion measures of proficiency in the aircraft. If they are not highly consistent for a given level of proficiency, no substantial correlation between them and a predictor measure is possible. Variations of flight conditions often result in fluctuations of measures that would be consistent for constant flight conditions Also, inflight measures often depend heavily on subjective judgments of persons obtaining the measures. Unless rating scales are clearly anchored to objective bases for judgments, criterion measures are likely to be low in reliability.

EXPERIMENTAL METHOD

Considerations for Experimental Control

Ideally, all candidate device cue and dynamic capabilities should be varied factorially in each experiment, for the need for a particular device capability could vary with the presence or absence of others. Generally, this approach would not be practical. Instead, all possibly significant design features should be identified and documented in the report, whether or not they are varied during the experiment. As a data base accumulated, it could be examined for patterns of device capabilities that affect prediction of aircraft performance.

The ATD capabilities under study would be varied from full to zero, or degraded, employment as applicable. Problems affecting certain variables, e.g., transport lag, would be avoided if previous research had shown them to be disruptive of performance. Tasks to be performed would be standardized as to conditions and requirements. Criterion aircraft proficiency measures would be obtained only when weather and other operating conditions conform to a predetermined set of requirements that can ensure reliable criterion data. In the event that performance under adverse conditions is at issue, such conditions would, of course, be chosen for aircraft proficiency measures.

Procedure

Subjects would be assigned randomly to all groups. Each group would perform in an ATD under a single condition of device cue or dynamic capacity present/absent (or degraded). All subjects would also perform the same skills in an aircraft under conditions normally used for proficiency assessments, but in this case standardized to the extent practical. Half of each group would "fly" the ATD first, and the other half the aircraft. Measures in both vehicles would be those commonly obtained during proficiency assessments, although special care would be taken to standardize the procedures and conditions for measures which typically vary in practice. Also, additional measures would be obtained during ATD performance of any variable that previous research--or well-grounded intuition--suggests might be predictive of aircraft performance.

SUBJECTS

Subjects would be aircrew personnel who normally would undergo the proficiency assessments of concern. However, it is crucial that poorly and marginally qualified aircrew personnel be included along with better qualified, so that both predictive outcomes cited earlier can be assessed. Over a series of studies, different experience levels would be represented, for the commonality of ATD and aircraft proficiency may well vary, for better or for worse for predictive purposes, with the robustness of skills that generally increases with experience. The number of subjects needed in each group would depend on the number of times the experiment is to be replicated under conditions as nearly identical as possible. Also because the criterion is improvement in prediction, "shrinkage" in multiple regression predictions can be [illegible] depending on the number of candidate variables employed. Thus, either [illegible] separate correlation per group must be reliable, or it must be cross-validated, in which case 60 subjects per group would be sufficient.

DATA COLLECTION AND ANALYSIS

Data would be collected as described above. Analyses would be correlational in some form, with the form depending on the nature of the data. The significance of the contribution of each ATD variable to prediction would be established by testing the difference between correlations of [illegible] measures and aircraft proficiency measures with and without [illegible] variability. Using data from a number of such studies, the prediction [illegible] from various combinations of device variables would be [illegible] through multiple regression analyses.

FACILITIES

A laboratory ATD, such as the ASPT, would be employed in the study. The ATD would be varied as appropriate during the study. The ATD would be used for a total of approximately 350 hours. It would be employed in Task 2 for 30 hours distributed over 3 months, and in Task 3 for 320 hours distributed over 7 months.

SCHEDULE

The study would take 18 calendar months and require 3.25 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

Task	Schedule (Contract Months) Task start	Task finish	Number of person-years
1. Selection of device factors for study.	1	2	0.25
2. Conduct of training and collection of data.	2	6	0.50
3. Assessment of the skills.	7	13	0.50
4. Data analysis.	11	16	1.00
5. Report preparation.	15	18	1.00
	Total person-years		3.25

15. DETERMINATION OF VISUAL CUE REQUIREMENTS

PROBLEM

Aircrew skill performance depends more on visual cues than any other kind of cue information. While intra-cockpit visual cues have been well provided for in most ATDs, representations of extra-cockpit cues are still in relatively rudimentary stages of development. Furthermore, as extra-cockpit visual cueing systems are presently being developed, the thrust is for realism of scene content and its representation. This thrust has not only resulted in expensive visual systems, but as some researchers have pointed out, realism of scene content may result in reduced training efficiency for some tasks.

The problem is that aircrews do not use scene content per se as cues. Rather, they extract cue information from scenes. Necessarily, the cues they learn to use transcend peculiarities of specific scenes. They comprise characteristics shared by any number of possible terrains, atmospheric conditions, other aircraft, etc. Futhermore, the cues used for any particular task vary with both the task and the kinds of cue information that can be gleaned from a given set of circumstances.

The question of desired scene content in ATD visual systems is complicated, because we can only guess at present what generalizable dimensions of scenes become cues to aircrews, and what alternative dimensions may be used as circumstances vary. Futhermore, even if the guesses were on target, we still would not know what cue dimensions are needed for training. Perhaps all are needed, but more likely a nucleus of dimensions would both permit adequate cueing for ATD training and provide a basis for the rapid learning of additional cues after transition to an aircraft. Such a basis for generalization and transfer is important, for, even with the best visual system technology, the scene representation will always be incomplete and imperfect.

The experiment outlined below addresses only the first part of this dilemma: What dimensions of scene content are used as cues by pilots as evidenced by their effects on skill performance, and under what kinds of circumstances? The tasks to be performed and the circumstances would be selected to represent a variety of possible visual cue requirements, thus increasing the likelihood that findings would have implications for other tasks and circumstances, as well as for the identification of a "nucleus" for training referred to earlier. However, additional research would be required to develop the full implications for ATD training.

RESEARCH OVERVIEW

The study would consist of the five major tasks shown below.

Task 1: Identification of a set of candidate visual cue dimensions and preparation of an interim report.

Task 2: Selection of flight tasks to be studied in terms of visual cues required for the performance of the tasks.

Task 3: Assessment of the performance of pilots who attempt to perform the tasks under varying visual cue conditions.

Task 4: Analysis of the data to determine which cues are used by pilots to perform each of the tasks.

Task 5: Preparation of a final report that describes (a) the analysis of visual cue dimensions, and (b) the results of the experiment assessing visual cue requirements for different tasks.

ANALYTIC REQUIREMENTS

The success of the experiment depends on how completely likely cue dimensions are identified. Thus, a list of candidate dimensions must be drawn up. Training literature has occasionally addressed this problem, and with considerable insight. However, suggestions from the literature directly concerning visual cue dimensions should be supplemented by analyses of training outcomes, with and without visual cueings, when dimensions of scene content were not themselves being deliberately studied. Perceptive analyses could discover patterns of visual cueing that, when viewed vis-à-vis well established principles of visual perception, reveal dimensions worthy of empirical study. Interviews with pilots would also be helpful, but they should be undertaken with a clear recognition that pilots seem often unaware of the cues they use.

When a list of candidate dimensions has been compiled, a set of flight tasks must then be selected for the experiment. These tasks generally must require, at least ostensibly, a variety of cues useful for a variety of purposes. The tasks chosen should be representative of the spectrum of aircraft visual tasks varying from simple aircraft attitude control to the interactive complexities of ACM and visual weapons tasks. Also, a number of different (visual) circumstances for task performance should be possible--day vs. night; high altitude vs. low altitude; cloudy vs. clear atmosphere; etc.

Finally, the list of candidate dimensions and tasks must be culled using six criteria: (1) the task requires extra-cockpit visual cueing; (2) the task can be performed in an ATD that will be used in the study, and in a manner that permits a realistic evaluation of performance; (3) a visual system is available that permits presentation of the cue in realistic (dimensional) form and permits withdrawing the cue entirely (it would also be desirable for the cue to be available in a degraded or

"weakened" form, so long as the cue is not distorted); (4) transport lags for cue onsets and changes are known not to be disruptive of performance; (5) duplications of task cue requirements are minimized except as needed to study cue preference and interactions; and (6) the final lists of cue dimensions and tasks are reasonable considering time and other constraints that must be satisfied.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The critical requirement is that some minimum of cue integrity be available in the visual display, and that the cue be removable (or degradable) without distorting other cue dimensions. It would probably be better if ostensible realism of scene content is kept at a minimum within limits of adequate cue representation so the extraction of cue information would not be either enhanced or interfered with by peculiarities of specific terrains.

Cue patterns must be varied systematically in ways that permit effects of cues, separately and jointly, to be identified. It is likely therefore that a considerable amount of effort would have to be devoted to developing and checking out software.

Procedure

Experienced pilots, with recently demonstrated aircraft proficiency in those tasks they are to perform, would fly the various tasks in the ATD(s). No pilot need fly all tasks, but patterns of task assignments would be devised so that analyses of correlated (within pilot) and uncorrelated (between pilots) performance would not be compromised by inextricable mixtures of the two classes of data. Also, all pilots would fly their tasks with complete cueing enough times to establish stable bases for comparisons when some cues are not available. These trials would occur at the beginning of a pilot's participation so that his familiarity with the device could be assured. A given task under a given set of conditions would be repeated enough times, either by the same pilot or by different ones, to establish stable mean results. Quality of performance would be measured as appropriate for each task, with emphasis upon objective indicators to the extent possible. Also, a carefully constructed questionnaire and/or interview following completion of reduced-cue trials would identify the pilots' subjective impressions of the cues, their adequacy, and their utilization of them.

SUBJECTS

All pilots used in the study should have recently demonstrated aircraft proficiency on all tasks they are to perform during the study.

The number of pilots needed depends upon the number of tasks to be flown and the number of cue patterns used per task. As a minimum, at least 2 pilots should perform each task under a given cue condition, with a third pilot used for those instances where disparate results are obtained for two pilots.

DATA COLLECTION AND ANALYSIS

Measures of quality of performance would be obtained on each trial of each task. Measures should be as objective as possible, and personnel responsible for obtaining the measures would be trained sufficiently to ensure accurate results. The analysis of the data would focus on relative quality of performance for the presence/absence of each cue separately, and in interactions with other cues. A complex univariate or multivariate analysis of variance design would most likely be used to identify cue and pattern effects. Follow-up analyses would be determined by results of the overall analyses.

FACILITIES

A laboratory ATD, such as the ASPT or SAAC, would be required for this study to enable precise control over manipulation of the content of the visual scene. The ATD that is selected would be employed in Task 3 for a total of approximately 110 hours distributed over 3 months.

SCHEDULE

The study would take 18 calendar months, assuming mathematical models will already exist for the aircraft to be simulated. Four and one-half person years of professional contract labor would be required. The calendar time and level of effort required for each task and the total study are as follows:

Schedule
(Contract Months)

Task	Task start	Task finish	Number of person-years
1. Identification of visual cue dimensions, and interim report preparation.	1	12	2.00
2. Selection of tasks to be studied.	11	12	0.25
3. Assessment of pilot performance under varying cue conditions.	13	15	0.75
4. Analysis of data.	14	16	0.50
5. Report preparation.	14	18	1.00
	Total person-years		4.50

16. SCENE CONTENT ALTERNATIVES FOR PRESENTING VISUAL CUES

PROBLEM

Specific details of scene content are of concern in ATD design because of the cost involved in providing various kinds and degrees of detail, and because of different effects various details and overall scene content may have on training. The issue is complicated by the fact that aircrews do not use scene content per se as cues. Rather, they extract cue information from scenes by identifying aspects or dimensions of the content that are relevant to the need at the time.

A separate research effort was proposed for identifying cue dimensions that pilots use, i.e., the invariant aspects of scenes that provide visual referents and guides for performing various tasks. A follow-on effort implied in the outline of that research would determine, in addition, which cue dimensions are needed specifically for training. The research proposed here addresses the issue of how the dimensions needed for training should be represented in scenes. It is one thing to know, for example, that the apparent converging of parallel lines is a distance cue. It is another thing to know whether it is best in training to represent such parallel lines as highways, rectangular fields, power lines, or even simple checker-board patterns. Further, distance cues can be presented through other content dimensions such as texture gradient, interposition, etc.

Almost no previous research has been concerned with this topic. Instead, degree of realism in content has been the focus. In a few studies, however, incidental observations revealed that the nature of scene content had a significant impact on pilot performance. In fact, simulated fields, laid out in realistically uneven contours, were found in one study to be relatively useless for cueing, while fields laid out unrealistically in patterns with only straight lines and right angles were very useful cues.

The problem appears to be that when viewing a real terrain, a pilot can scan it for whatever features he can find to use as cues; but in a simulated CGI terrain, for example, he can use only those actually programmed into the computer that generates the image. Hence, it is important to know what to put into the computer or other image generating mechanism.

There are countless variations in possible scene contents. Not only may given cues be represented in a variety of scenic details, but the nature of the task dictates classes of representations as well as degrees of realism. A number of referents might be used to judge motion relative to the ground, but if an object on the ground, say, a tank, is to be a target for a bomb drop, it must be identifiable as a tank. Even so, it would not necessarily have to resemble a tank, unless it is to be

distinguished from some other similarly tank-shaped or -sized object that is a nontank. If the objective of training is to teach real-world target recognition and identification of tanks, the realism of the display content does, of course, become of concern.

The research outlined below is a prototype experiment for determining the effects of different scene cue representations on skill acquisition in ATDs, and on transfer of ATD-trained skills to aircraft. A number of replications of the experiment will be needed to study the most efficacious representations of cues for separate tasks. Also, understanding of training values of various cues, which is very limited at present, will increase if this research is carried out, visual simulation technology will improve, and a data base regarding effects of cue representations will accumulate. These three advances will provide a continuing impetus to refinements in scene contents.

RESEARCH OVERVIEW

The study would consist of the five major tasks shown below.

Task 1: Selection of tasks to be investigated.

Task 2: Identification of candidate characteristics of visual scene content believed to be related to the effectiveness with which each task is learned.

Task 3: Development of ATD training and of measures for assessing trainee performance.

Task 4: Comparison of the effectiveness and efficiency of performance of trainees who are trained in the aircraft with that of students who are trained under varying visual scene content conditions.

Task 5: Analysis of data and preparation of final report.

ANALYTIC REQUIREMENTS

A necessary condition for the success of any experiment on effects of various representations of visual cues is that the nature of the cues needed be known. Presumably, other research would have established types of cue information needed for training performance of various tasks. This research should also be examined carefully for suggestions as to effective representations of cues. In addition, any training research using extra-cockpit visual scenes should be reviewed for suggestions, and literature on visual system technology should be consulted so that decisions regarding the design of scenic details would be consistent with technological capability. In the earlier experiments, existing

programs for visual displays could probably be adapted for experimental purposes with relatively little difficulty.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The efficacy of given visual scene representations will vary as cue requirements differ for separate tasks. The concern is scene requirements for separate tasks, and for those groups of tasks that have similar scene requirements. Therefore, tasks to be trained during the experiment should be homogeneous in visual cue requirements. However, even though these requirements are ostensibly the same, amount of detail of object characteristics, and interactive effects of cues within patterns, could still result in different efficacies of given scenes for different tasks. Hence, the experimental design should permit collection of performance data separately for each task.

Objects in simulated scenes would not be appreciably disparate from the terrain where aircraft transfer trials will be flown. (It is not necessary to duplicate the actual terrain; transfer, i.e., generalization, of visual cueing is the object.) However, realism in representations would be varied.

The instructional regimen would be standardized for all students. Instructors would be trained to follow the standard instructional procedures.

Procedure

Subjects would be assigned randomly to all groups. A single control group would either receive ATD training without the visual display (if practical), or receive all training in an aircraft. (If the latter, subsequent aircraft trials for all groups should be in a different terrain area so that generalization will be required of the control group as well as ATD-trained groups.) All other groups would be trained in the ATD, with a different set of visual cue representations for each group. Level of performance would be measured periodically during training, and at the end of ATD training (except for an aircraft control group). ATD subjects would transition to an aircraft immediately upon completing ATD training. Their initial performance level in the aircraft would be measured, and additional measures would be made periodically until proficiency was achieved.

SUBJECTS

For whatever tasks trained, all subjects should have comparable, and usual, previous experience. There should be multiple subjects in each

group, with the number per group sufficient to provide at least 18 inter-subject degrees of freedom in the error term for all comparisons.

DATA COLLECTION AND ANALYSIS

The periodicity of measures would be adequate to reveal the rate of skill acquisition in the ATD and rate of progress to proficiency in the aircraft. Measures would also be obtained in the ATD at the completion of ATD training, and on the first or early trials in the aircraft.

The analyses would focus on rates and levels of achievement as revealed by these measures. Special attention would be given to ascertaining whether negative transfer to aircraft occurs as evidenced, separately, by original aircraft performance or rate of progress to proficiency. Multiple group, repeated measures analyses of variance would be used with training tasks comprising the repeated measures.

FACILITIES

A laboratory ATD in which visual scene content can be manipulated would be employed in the study. It would be used for a total of approximately 100 hours. In Task 3, it would be employed for a total of approximately 20 hours distributed over 3 months, and, in Task 4, for a total of approximately 80 hours distributed over 4 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

It is assumed that this effort would build on the results of the previous study of visual cue requirements. Given that background, the present study would take approximately 16 calendar months to conduct and 3.00 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

Schedule
(Contract Months)

Task	Task start	Task finish	Number of person-years
1. Selection of tasks to be studied.	1	2	0.20
2. Identification of candidate characteristics of visual scene content.	1	6	1.00
3. Development of training and measures.	5	8	0.30
4. Conduct of training and collection of data.	9	12	0.50
5. Data analysis and report preparation.	11	16	1.00
	Total person-years		3.00

17. DETERMINATION OF MOTION AND FORCE CUE REQUIREMENTS

PROBLEM

Patterns are emerging in results of research on force and motion cueing in ATDs. Visual cues are obviously needed for some tasks; g-cues arising from g-seats, g-suits, and helmet and arm loaders have facilitated performance in a few studies; and platform motion appears to be of little use for training certain tasks. In addition, seat and stick shakers have been studied, but not in a context where their value for training could be assessed.

Visual cue requirements are addressed in two other research efforts proposed in this report. The concern here is with the analysis of motion and force cue requirements that may be provided by other than visual means.

The motion and force cues used by aircrews have not been systematically identified. It is not known which particular sensory mechanisms are involved or how they may be stimulated most effectively for training purposes. Most research efforts in this area have centered on platform motion, and while, as noted, such motion has been found of little training value in several studies of specific tasks, there remains some lack of agreement on motion. For example, two reviews of research on platform motion concluded that platform motion may be useful when it alerts the pilot to forces arising outside the control loop. One reason for the lack of agreement in this area stems from the fact that most of the research has been based on a stimulus fidelity approach rather than a cue learning approach.

The purpose of the research outlined here is to identify nonvisual motion and force cues aircrews use during performance of various skills. The orientation is a task-behavioral one rather than stimulus fidelity. From this knowledge, follow-on efforts as described in a subsequent outline of an experiment could determine which of these cues are needed for training in ATDs, and how they can best be represented.

RESEARCH OVERVIEW

The study would consist of the four major tasks shown below.

Task 1: Development of a set of candidate dimensions of motion and force cues, and preparation of an interim report.

Task 2: Selection of tasks to be studied in terms of motion and force cueing requirements.

Task 3: Performance of the tasks by subjects in the ATD under varying conditions of motion and force cueing.

Task 4: Analysis of performance data to determine the separate, joint, and redundant effects of motion-force cueing on performance of the tasks, and preparation of the final report.

ANALYTIC REQUIREMENTS

To a great extent, this would be an analytic effort devoted to syntheses of research literature focusing on motion and force cueing. A number of studies have used different cueing mechanisms that provided different kinds of information on the one hand, and through different body receptors on the other. In reviewing this research, the focus would not be on one cueing device as opposed to another, but on identifying sensory inputs arising from devices that appear facilitative of performance and learning versus inputs from devices that do not aid performance.

In addition, the psychological literature on perception of motion and force would be examined thoroughly. While occasional training studies have alluded to well known principles of visual perception in analyses of visual cues, comparable uses of psychological knowledge regarding motion and force perception have been much less frequent. Instead, attempts have focused on duplicating motion and gravity effects simply by providing quasi motion and quasi accelerative conditions. With this emphasis on device characteristics as opposed to perceptual principles, there has been little concern with understanding how force and motion cues are integrated kinesthetically to coordinate motor and cognitive actions and how such factors may relate differentially to various tasks.

The product of the analytic efforts would be a list of candidate dimensions of force and motion cueing. To the extent possible, the integrative use of each dimension in different classes of task performance would be identified as well. Experiments would then be run to determine which cues affect performance of experienced aircrews.

EXPERIMENTAL METHOD

Considerations for Experimental Control

Candidate dimensions would be held constant, and manipulated, singly and in patterns that permitted the assessments of (1) separate contributions of each dimension; (2) joint contributions of combinations of dimensions; and (3) redundancy of cue information. Each perceptual dimension would be represented by at least one available device, with alternative devices used when practical. Cueing would be both of an either/or nature, and with partial degrading when possible without

distortion. Tasks to be performed would be standardized, and would represent as far as practical a broad range of aircrew skill performance. All subjects would have recently demonstrated aircraft proficiency on each task.

All personnel involved in running the experiment would be trained to fulfill their roles in a standard manner.

Procedure

Subjects would be assigned randomly to groups. Each subject would be given preliminary practice in the ATD sufficient to familiarize him with the device. Each group would "fly" the ATD for at least one factorial combination of dimensions/device representations. Preferably, the same subjects would "fly" under more factorial combinations, even all of them, unless too many are involved, or successive trials would lead to criterion contamination due to fatigue or practice effects. Also, aircraft practice would be systematically interspersed with ATD trials to avoid over-adaptation to the limitations of the device. Counter-balancing of conditions and aircraft trials would avoid otherwise unidentifiable cumulative effects.

Measures would be obtained on each trial as appropriate to assess each task performance, and in sufficient detail to reveal separate aspects of performance. Subjects would also be queried via questionnaires and interviews regarding their subjective evaluations of the cues available and felt need for additional cues.

SUBJECTS

All subjects should be aircrew personnel who had recently demonstrated aircraft proficiency on the tasks used. Total numbers of subjects would vary with the number of experimental conditions used for each subject, but at least 6 subjects should be used for any one factorial condition.

DATA COLLECTION AND ANALYSIS

Measures would be obtained on each trial as appropriate for each task, using standardized procedures and techniques. Data analyses would focus on similarities and differences in performance that revealed separate, joint, and redundant effects of cueing on performance. A complex analysis of variance would be used for most data, with the correlated main effects of multiple tasks, and of multiple conditions for separate groups, represented in a way that correlated and uncorrelated measures would not become mixed in the analyses.

FACILITIES

The experiment would be conducted in a laboratory ATD in which motion-force cues could be presented and varied through a number of different motion-force cueing devices. The ATD would be employed in Task 3 for a total of approximately 110 hours distributed over 4 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

Given the existence of an appropriate ATD, the study would take 16 calendar months and require 3.00 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

	Schedule (Contract Months)		
Task	Task start	Task finish	Number of person-years
1. Development of candidate motion and force cues, and *interim report preparation*.	1	9	1.50
2. Selection of tasks to be studied.	9	10	0.20
3. Performance of the tasks under varying motion and force cueing conditions.	11	13	0.50
4. Data analysis and report preparation.	12	16	0.80
	Total person-years		3.00

18. ALTERNATIVE MECHANISMS FOR PROVIDING MOTION-FORCE CUEING

PROBLEM

While some patterns are emerging in results of research on motion-force cueing in ATDs, there is no consensus on the need for such cues or on the means of providing them. G-cueing through simulated stimuli arising from g-seats, g-suits, and helmet and arm loaders has been shown in some studies to aid coordination of control movements. Stick and seat shakers have less empirical support as effective mechanisms, but at the same time they have not been demonstrated to be ineffective. And there is a general impression among some training researchers that platform motion makes little or no contribution to aircrew training for certain skills.

However, significant questions remain, even regarding platform motion. Studies of the effects of platform motion on transfer of ATD-trained skills to aircraft, while sometimes well conceived, have often used statistical analyses inappropriate for the data, and valid interpretations are impossible. As for most other g-cueing devices and stick and seat shakers, their effects on transfer are not known. Futhermore, in studies showing improvement in performance in ATDs, experienced pilots have typically served as subjects. Thus, it also has not been established that these cueing mechanisms even aid skill acquisition in ATDs.

The research outlined below recognizes the complexity of the effort needed to establish clearly the value of various motion-force cueing mechanisms in training. The experiment is a prototype that may be conducted at various levels of complexity, depending on the resources available for a particular study. Replications of the experiment, using various factorial combinations of cueing devices, will be necessary to sort out all the separate effects and their interactions and redundancy (for training).

The question to be answered is, What are the separate and interactive effects of the force and motion cueing devices used (1) upon skill acquisition in ATDs and (2) upon transfer of ATD-trained skills to aircraft?

RESEARCH OVERVIEW

The study would consist of the four major tasks shown below.

Task 1: Selection of the tasks and motion-force cueing mechanisms to be studied.

Task 2: Development of ATD training and of measures of trainee performance.

Task 3: Assessment of the differences in performance of students trained under varying motion-force cueing conditions.

Task 4: Analysis of the contribution of each motion-force cueing mechanism to training effectiveness and efficiency for each of the tasks studied, and preparation of the final report.

ANALYTIC REQUIREMENTS

No prior analytic efforts would be needed other than those directly involved in implementing the experiment and instrumenting measures.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The manipulation of cueing mechanisms under study should be such that three kinds of information are obtained regarding each: (1) the effect of each mechanism separately; (2) the redundancy of each pair of mechanisms; and (3) the joint effects of each pair (or triple, etc) of mechanisms. Tasks to be trained would vary as much as practical so as to represent a wide range of skills, thereby permitting broader generalizations of findings. The training regimen would be standardized, and instructors trained to implement it in a standard manner.

Procedure

Subjects would be assigned randomly to each experimental group and to a control group that received no force-motion cueing of any kind. Some separate experimental groups would receive (only) the cueing from each source. Other separate groups would receive (only) one combination of a single pair of types of cues; still other separate groups a single triple of types of cues, etc. Each subject would be trained on all tasks used. Objective measures as well as instructor ratings of performance would be obtained periodically during and at the end of ATD training.

Immediately following ATD training all subjects would transition to an aircraft and continue practice to proficiency on the ATD-trained skills. Measures of performance would be obtained on the first, or a set of early trials, in the aircraft and periodically until proficiency is achieved.

SUBJECTS

The subjects should be pilots at a comparable level of training and experience. (The level should vary often enough from one experiment to another so that different levels of competence and of task difficulty would eventually be represented.) No subject would have had prior training on the skills to be taught during the experiment. The number of subjects per subgroup should be adequate to provide at least 18 between-subject degrees of freedom for each significance test.

DATA COLLECTION AND ANALYSIS

A necessary minimum of instructors would be trained to obtain measures in a standardized manner. If preliminary reliability checks on any measure reveal inter-instructor reliabilities below .75, the average of measures obtained by at least two instructors would be used in analyses involving those measures.

The periodicity of measures would be adequate to reveal rate of skill acquisition in the ATD and rate of progress to proficiency in the aircraft. Measures would also be obtained in the ATD at the completion of ATD training, and on the first or early trials in the aircraft.

Summaries of data (e.g., means, standard deviations, percents) would be compiled as appropriate. Multiple group repeated measures analyses of variance would be used with training tasks comprising the repeated measures. (Depending upon types of groups and adequacy of overall experimental controls, the various groupings according to cueing patterns could be combined into a single complex analysis of variance.)

FACILITIES

The study would require an ATD with all available motion-force cueing devices. A laboratory ATD should be employed to enable precise control over use of mechanisms. The ATD that is selected for this study would be used for a total of approximately 215 hours. It would be employed in Task 2 for a total of approximately 25 hours distributed over 2 months, and in Task 3 for a total of approximately 190 hours distributed over 4 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

Given the existence of an appropriate ATD, the study would take 16 calendar months and 3.00 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

Schedule
(Contract Months)

Task	Task start	Task finish	Number of person-years
1. Selection of tasks and motion-force cueing mechanisms to be studied.	1	3	0.50
2. Development of training and measures,	4	7	0.50
3. Conduct of training and collection of data.	8	11	0.75
4. Data analysis and report preparation.	11	16	1.25
	Total person-years		3.00

19. ALTERNATIVE STRATEGIES FOR PRESENTING EXTERNAL VISUAL SCENES TO THE INSTRUCTOR

PROBLEM

Two issues are involved here. First, the optimum display strategy to present external visual scenes to instructors depends on what the instructor needs to know. This need will vary with the tasks being practiced by students, and probably with the level of the student in some instances. Second, the optimum strategy varies as alternative modes of presentation facilitate or hinder the instructor's processing of the information they contain. Extracting information from multiple displays of varying analytic detail, while simultaneously collating the information to direct instruction, can be a harrying task.

For a number of learning tasks, an X-Y plot can be both the simplest to present, and the most informative if an altitude reading is available when needed. The concern here, however, is the instructor's opportunity to recognize what a student responds to, and should respond to, as visual cues. An obvious starting point is to let the instructor see duplicates of everything seen by the student. There are two disadvantages to this approach. First, visual displays used by students are usually quite expensive. Similar capabilities at the instructor's console could involve unnecessary expense. Second, the purpose of the student's visual scene is to provide cues for learning and for task performance. The purpose of the instructor's display is to provide information needed to <u>teach</u> task performance. His display may well need to incorporate less, <u>or more</u>, information of various types than does the student's. In some instances, the instructor will need information for diagnostic purposes that would never appear as such in a "real world" scene. Conversely, the student may require types and levels of visual information in the learning of discriminations that are of no utility to the instructor for his instructing functions.

The prior question in this case is the analytic one: What does an instructor need to know? Specifically, what information is needed to assess the status of the student's cue and response discriminations relative to task requirements so that guidance and feedback can be provided accordingly? When this question is answered, task by task, candidate alternatives for providing the information can be identified for assessment through experimental analyses.

Nevertheless, the analytic and empirical issues can be pursued simultaneously. Available alternatives for instructor displays can be evaluated for training various tasks even while only partial answers are known--or hypothesized--to the analytic question. Generally, the display alternatives may be grouped as: (1) a full-scale reproduction of what the student sees; (2) a reduced-scale reproduction of what the student sees; (3) reproduction of only selected aspects of what the student sees;

(4) only graphic representations of what the student sees; and (5) combinations of the first or second with the third and/or fourth, or of only the third and fourth. In any case, experimental evaluations should focus on alternatives that can reasonably be expected to provide information useful for instruction as determined by analyses of the learning processes involved and of an instructor's ability to process the information with the training and experience he will have had.

A programmatic research effort is entailed, for which a prototype experiment is outlined below. The question to be answered is, Which type of available display (or displays to be designed following analytic development) can an instructor use most efficiently and effectively for teaching given tasks?

RESEARCH OVERVIEW

The study would consist of the four major tasks shown below.

Task 1: Identification of information that contributes to the capability of the instructor to assess the status of the student's cue and response discriminations.

Task 2: Development and conduct of instructor training.

Task 3: Use of standard test scenarios to assess instructor performance under varying external scene display conditions.

Task 4: Analysis of the data and preparation of the final report.

ANALYTIC REQUIREMENTS

The general nature of the analytic requirements was identified above. Specifically, a partial learning analysis would be needed, task by task, to identify (1) discriminations of cues and responses to be learned by the student; (2) cues/responses/conditions most likely to introduce interference with the learning desired; and (3) what an instructor needs to know to assess the status of a student's cue and response discriminations. Thorough reviews of the literature would be necessary, as well as careful logical analyses that focus on how aircrews process cue information.

EXPERIMENTAL METHOD

Considerations for Experimental Control

The essential requirements for experimental control are that tasks be clearly defined, conditions for their performance be standardized, and

indicators of needed specific instructor participation occur in the visual displays. The training regimen and contingencies for instructor actions of concern should be standardized as well. Although actual students could be used to perform tasks, it would be much better if instructors, fully trained to exhibit anticipated difficulties in skill acquisition, did so in standard scenarios. (Preprogrammed scenarios would be even better.) Subject instructors would have received whatever prior preparation and training they normally would have for teaching the skills in the ATD.

Conditions for obtaining all dependent measures would be standardized, and all personnel involved in implementing the experiment would be trained to fulfill their roles in a standard manner.

Procedure

Instructors who serve as subjects would be assigned randomly to groups. Each group would have information displayed according to one of the five possibilities given earlier. (With proper counterbalancing and appropriate choices of statistical analyses, each group could use more than one type of display.) Scenarios would be presented, and measures would be obtained that revealed (1) the recognition, or lack thereof, of indications of need for specific instructor intervention; (2) the appropriateness of the intervention (if any); and (3) the appropriateness and acceptability of the delay in the instructor's reaction.

SUBJECTS

Subjects should be instructors qualified to teach the skills at issue in the ATDs used. Numbers of groups could vary with each experiment, but the number of subjects per group should be adequate to provide at least 18 inter-subject degrees of freedom for any group comparison.

DATA COLLECTION AND ANALYSIS

Data would be obtained as indicated above. Analyses would focus on adequacy of display alternatives for given tasks. Thus, a *t* test or one-way analysis of variance would be appropriate, extended for repeated measures as needed to incorporate multiple observations on dependent variables. If more than a single display format is used with each group, an additional (correlated) dimension would be added to the basic variance analysis.

FACILITIES

An ATD should be used that permitted all display options to be employed. For full scale reproduction of the visual scene, slaved

cockpits could be used. The remaining options would be available at the IOS. The ATD that is selected for this study would be employed for a total of approximately 220 hours. It would be employed in Task 2 for a total of approximately 100 hours distributed over 2 months, and in Task 3 for a total of approximately 120 hours distributed over 4 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

Given the availability of an appropriate ATD, the study would take 15 calendar months and 2.50 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

	Schedule (Contract Months)		
Task	Task start	Task finish	Number of person-years
1. Identification of information that contributes to the capability of the instructor to assess the status of the student's discriminations.	1	8	1.25
2. Development and conduct of training.	8	9	0.25
3. Assessment of instructor performance under varying external scene display condtions.	10	13	0.25
4. Data analysis and report preparation.	11	15	0.75
	Total person-years		2.50

20. EFFECTS OF INSTRUCTOR LOCATION ON TRAINING IN VISUAL SIMULATORS

PROBLEM

The location of the instructor in future visually equipped simulators, either on board or at a remote console, will affect simulator design and therefore cost. His location might also have consequential effects on training. It is important, then, to determine whether training would be facilitated, or adversely affected, by one location more than the other, or whether instructor location is not in itself a training effectiveness issue.

The experiment proposed here would examine the effects of on-board vs. remote location of instructors on training. Air combat maneuvers (ACM) would be taught in the study. ACM skills are selected because they require visual information for orientation, cueing, and execution, as well as continuous visual feedback regarding appropriateness of control inputs. In addition, several previous investigators have reported sharply contrasting results as to the transfer of simulator-trained ACM skills to aircraft performance. One possible reason that has been advanced to explain this discrepancy in transfer results is the different locations of instructors during the simulator training. Therefore, a focus on ACM training not only ensures that visual capabilities will be exploited, but offers an opportunity to resolve the issue of instructor location in a training problem of considerable significance.

The essential question to be answered by the proposed experiment is: Do on-board, as opposed to remote, locations of the instructor during simulator training have differential effects on (1) rate and level of ACM skill acquisition in the simulator, and (2) on the amount and quality of skill transfer to aircraft performance?

RESEARCH OVERVIEW

The study would consist of the three major tasks shown below.

Task 1: Development of (a) ATD training models appropriate to each condition of instructor location, and (b) measures of trainee performance.

Task 2: Assessment of differences in the performance of students trained by instructors on board the ATD in comparison with that of students whose instructors are at a remote location.

Task 3: Analysis of the data and preparation of the final report.

ANALYTIC REQUIREMENTS

No prior analytic efforts would be needed other than those directly involved in implementing the experiment and instrumenting measures.

EXPERIMENTAL METHOD

Considerations for Experimental Control

Instructor location can affect quality of simulator training only as it facilitates or constrains interactions with students. The nature and quality of these interactions depend directly on (1) the information available to the instructor; (2) his ability to interpret the information; and (3) his knowledge of effective and efficient uses of feedback and guidance. The training of instructors to be used in the experiment should ensure (2) and (3). As for the information available to the instructor, whether at the cockpit or remote console the instructor should receive whatever information would be provided at that location during operational use of the simulator. The intent is not to standardize the information received at each location, but to equate it, separately by location, with operational reality.

The nature and quality of instructor-student interactions also depend on indirect influences arising from their contacts. Briefly, instructors and students who "know" each other communicate more fully than those who do not. The disembodied voice of a familiar instructor over earphones has more personal meaning than that of a stranger; and insofar as emotional support (or threat) is concerned, the presence of a familiar monitoring instructor at a remote console may well be psychologically equivalent to his presence in a cockpit. Thus, the effects of instructor location on training may well vary with whether a student has the same instructor for all simulator sessions or a different one from day to day.

Whether the same or different instructors are used for a given student during the experiment should be determined by operational practices during training, so that effects of instructor location could be related to operational conditions. If research resources permit, of course, it would be better to use the same instructors for one part of the experiment and different ones for another part, thus revealing the interactions (if any) of this variable with instructor location. With this added comparison, results could be generalized more readily across training programs that differ in this particular regard.

Procedure

For the basic comparison, students would be assigned randomly to two groups. (These basic two groups would be randomly subdivided as necessary to accommodate patterns of instructor usage as above or any

the systematic variation in additional independent variables.) Both groups would undergo the same training regimen, but each member of one group would have an instructor in the cockpit with him (or at a jumpseat by the student station), while instructors for the second group would be at a remote console.

For trial engagements in the aircraft, specifically trained instructors would fly adversary aircraft according to a standard plan for trial engagements, devised specifically for the experiment. Instructors rather than other students would be used for this purpose so as to permit relative standardization of adversary actions. Also, safety would be less of a problem, and all adversaries would have a comparable level of competence.

SUBJECTS

A minimum of 20 subjects, 10 per basic group, should be used for the basic comparison. If basic groups are further subdivided, the total number of subjects should be adequate to provide at least 18 inter-subject degrees of freedom for any comparison. All subjects should be drawn from a pool of students typical of those who would undergo for the first time the training provided during the experiment.

DATA COLLECTION AND ANALYSIS

A necessary minimum of instructors would be trained to obtain the measures in a standardized manner. If preliminary reliability checks on any measure reveal inter-instructor reliabilities below .75, the average of measures obtained by at least two instructors would be used for analyses involving those measures.

Measures on dependent variables would be obtained at periodic intervals during skill acquisition in the simulator; at the completion of simulator training; on the first trial, or for each of a set of early trials, in the aircraft; and at appropriate later times in the aircraft. Actual schedules for data collection would depend on skills being taught and expected rates of progress. The schedule would be such as to permit comparisons of (1) rates of skill acquisition in the device; (2) levels of skill acquisition on completion of device training; (3) levels of performance at the outset in the aircraft; and (4) rates of progress to proficiency in the aircraft.

Summaries of data (e.g., means, standard deviations, percents) would be compiled as appropriate. The statistical significance of group differences would be determined by appropriate analyses (e.g., repeated measures analysis of variance for multiple groups) to reveal nonchance differences in rates of progress as in (1) and (4) above, and levels of achievement as in (2) and (3).

FACILITIES

Facilities requirements are driven by the need for a dome projection visual system (real imagery rather than virtual imagery) and the need for provisions to locate the instructor either in the cockpit (or at a jumpseat) as well as locate him at a remote console. Candidate facilities are: (1) the Northrop LAW/WAVS; (2) the NASA Langley DMS and (3) the McDonnel Douglas MACS. The simulator that is selected for this study would be employed for a total of approximately 220 hours. It would be employed in Task 1 for a total of approximately 20 hours distributed over 3 months, and in Task 2 for a total of approximately 200 hours distributed over 9 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

The study would take 15 calendar months to complete and require 3.00 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

Task	Schedule (Contract Months) Task start	Task finish	Number of person-years
1. Development of training models and performance measures.	1	3	0.75
2. Training and evaluation of student performance.	4	12	0.75
3. Data analysis and report preparation.	10	15	1.50
	Total person-years		3.00

21. REQUIREMENTS FOR DELAYED REMOTE DISPLAYS OF PERFORMANCE

PROBLEM

Most new ATDs offer excellent oportunities for feedback regarding adequacy of performance during practice sessions via record/replay features, and during debriefings following sessions via hardcopy printouts. Distinct advantages of record/replay during sessions are the opportunities to isolate separate aspects of performance for scrutiny, feedback, and guidance while the cues and actions used by the student are still freshly in mind. Distinct advantages of hardcopy printouts are the opportunities they provide for later analytic, meditative--mediational--contemplation of what a student did, why he did it, and how performance could be improved.

Use of hardcopy printouts during debriefing is typically hampered by the inability of the student, and often the instructor, to identify precisely what the conditions and step-by-step actions were at the time a decision was made or an action was initiated, altered, or terminated. A replay during debriefing of the actions would probably aid the mediational analysis and synthesis considerably. However, the value of delayed remote displays of performance segments surely varies with the training task at hand and with the experience of the student. In either case, the training issue reduces to the type and amount of information needed in the playback and how it should be presented.

The prototype experiment outlined below addreses these three questions: (1) What kind of information is needed? (2) In how much detail? (3) How should it be presented? A programmatic effort will be required, but fortunately a lot can be learned through initial experiments with current record/replay features of ATDs.

RESEARCH OVERVIEW

The study would consist of the three major tasks shown below.

Task 1: Development of a regimen for ATD training that incorporates and exploits the use of a delayed remote display.

Task 2: Comparison of the effectiveness and efficiency of performance of students who are trained using the delayed remote display during debriefs with that of students who do not have access to the display.

Task 3: Analysis of the data and preparation of the final report.

ANALYTIC REQUIREMENTS

The prototype experiment would require no prior analytic effort other than those necessary to implement the experiment and instrument measures. Once a data base accumulates, critical examinations of research results would, of course, be necessary for decisions regarding the design of delayed remote displays.

EXPERIMENTAL METHOD

Considerations for Experimental Control

For delayed remote displays to be of value, they should have demonstrable utility for efficient and/or effective training as evidenced by (1) improved skill acquisition in an ATD; (2) improved cognitive comprehension of requirements and contingencies for skill performance; and/or (3) increased transfer of ATD-trained skills to aircraft performance. Thus, all extraneous variables that could affect these dependent variables should be held constant. Instructor utilization of experimentally manipulated information should be standardized as to strategy, but allowed to vary in content as needed to focus on detail of individual student performances. The instructors, of course, should be comparable in knowledge of instructional technology. Even so, the same instructors should be used for all experimental groups, or else enough instructors should be available so that random assignments to groups would remove any systematic instructor effects. The training regimen would be the same for all groups except for the utilization of delayed replay.

All measures would be obtained under standardized conditions, using specially trained personnel.

Procedure

Subjects would be assigned randomly to (usually) two groups. Each group would be trained per the standard regimen. One group would undergo standard debriefings using whatever aids (e.g., hardcopy printouts) are usually available. The other group would also have these aids as needed, but in addition, at least part of their debriefings would occur at the remote terminal where previously recorded segments of their performance could be played back as desired. Verbal tests of cognitive comprehension of specific cue/response task requirements would be administered at appropriate times, and measures would be obtained during ATD practice to reveal rate of skill acquisition, at the end of ATD practice to show level of achievement, at the beginning of aircraft practice to show original level of proficiency, and during aircraft practice to reveal rate of progress to proficiency

SUBJECTS

Subjects should be aircrew students who normally would undergo training in the tasks used in the experiment. Program wide, a variety of tasks, and levels of tasks, would be used so that the value of delayed remote displays could be related to levels of student proficiency. At least 10 subjects should be in each group.

DATA COLLECTION AND ANALYSIS

Measures would be obtained as above so as to show rate of skill acquisition in the ATD, level of proficiency at the end of ATD training, original proficiency in the aircraft, and rate of progress to proficiency in the aircraft. Cognitive measures would be obtained at such times as needed to show the relation between level of understanding of task requirements and use/non-use of delayed remote playback.

In addition, students and instructors for the delayed display groups would be queried by questionnaires and interviews regarding the features, aspects, details, etc., they found useful, of no use, and of questionable value. They also would be asked to suggest other information they could have used but which was not available.

Basic analyses of data would compare the two groups on all measures, either as t tests or repeated measures analysis of variance as appropriate. Intercorrelations of cognitive and performance measures would be determined. Questionnaire/interview data would be cross-classified so as to reveal felt needs for informational details according to tasks and types of problems.

FACILITIES

A remote computer-driven video terminal would be the best instrumentation for the study. For the present effort, however, IOS displays could be employed (but it should be noted that employment of remote terminals in the future would reduce such demand for the ATD and IOS). The ATD that is selected for the study would be employed for a total of approximately 170 hours. It would be employed in Task 1 for a total of approximately 20 hours distributed over 2 months, and in Task 2 for a total of approximately 150 hours distributed over 9 months.

SCHEDULE AND CONTRACTOR PERSONNEL REQUIREMENTS

The study would take 15 calendar months to conduct and require 2.50 person-years of professional contract labor. The calendar time and level of effort required for each task and the total study are as follows:

Schedule
(Contract Months)

Task	Task start	Task finish	Number of person-years
1. Development of training that exploits use of a delayed remote display.	1	4	0.75
2. Conduct of training and collection of data.	5	13	0.75
3. Data analysis and report preparation.	11	15	1.00
	Total person-years		2.50

APPENDIX C

BRIEF RESEARCH PLANS

The topics included in this appendix were judged to be important with respect to their potential impact on ATD training, but of lesser priority than the topics in Appendix B. Brief research plans are provided that discuss the nature of each research topic and outline certain analytic and experimental issues of concern.

The topic titles are shown below. Page numbers are included to aid in locating each plan.

1. GUIDES FOR LEARNING ANALYSES

ISD and similar approaches to training problem analyses have had a valuable impact on military training. However, for optimum use in program design, it is necessary that all cues involved in skills be already known, and alternative cues needed for training be clearly identified. Just as crucial is the necessity for personnel making the analyses to use considerable intuitive expertise to design training scenarios that exploit all aspects of efficient and effective learning.

These conditions are rarely met in practice, and the designs of ATD training programs reveal that these prerequisites are not fully understood. Physical correspondence of device characteristics to the aircraft and flight environments substitutes for a knowledge of cues actually used by aircrews; cues used in training are determined by compromises regarding device technology and cost; and the complexities of the transfer of ATD training to aircraft are glossed over by equating transferability of ATD learning with device and task fidelity.

Learning analyses would go beyond ISD by determining, for example, (1) cues used by aircrews; (2) cues needed for training; (3) student capabilities to substitute mediation for physical cues; (4) optimum conditions (total cue complex, practice schedules, employment of guidance and feedback) for developing stable cue discriminations, separately and as integrated into cue-response patterns; (5) optimum conditions for teaching of cue-response discriminations that would be resistant to interference and forgetting; (6) transferable dimensions of cues and responses so as to maximize transfer; (7) cognitive bases of skill performance; (8) sequencing of learning to maximize contributions of cognitive and noncognitive mediation; (9) experiential requirements for effective mediational development; and (10) alternatives available for adaptation of any aspects of this training as indicated by student, device, aircraft, and performance characteristics.

Aircrew skills are complex; and when taught in ATDs, the complexity of teaching them is compounded by dependence on transfer, a complicated process in itself. It is because of this complexity that learning analyses are needed. No existing procedure for developing training programs addresses any of the example topics in any depth, if at all. In fact, there is no extant systematic guidance for attacking such issues. In brief, learning analyses would pick up where ISD leaves off.

The research proposed here would develop a systematic format and procedural guide for performing learning analyses. The focus would be on ATD training, but the format and procedures would be generalizable to any kind of training.

2. RELATIONSHIPS BETWEEN CONDITIONS FOR PRACTICE ON SEPARATE TASKS AND THE SUBSEQUENT INTEGRATION OF THE TASKS

The way in which tasks are organized for practice can influence the efficiency and effectiveness of skill acquisition. Learning is often facilitated during the early stages of skill acquisition by the separation of tasks for practice in ATDs that are normally performed concurrently or in succession in the aircraft. As these separated skills are developed in ATDs, the student should then be taught to integrate these separately practiced tasks so that he will be prepared to perform in complex operational situations.

Such integration will eventually occur simply if the larger task is practiced as a whole after separate parts have been learned. As discussed in the Utilization report, however, the integration of separately learned tasks will be made easier if provisions for their integration are made when the separate tasks are practiced. Unfortunately, knowledge concerning such provisions has not been formulated into applied principles that could be used by nonspecialist military personnel who frequently design ATD training programs. The guidance provided by ISD procedures, for example, is not sufficiently specific to enable many task level decisions to be made by these personnel.

As part of programmatic investigations of learning analysis procedures, research is needed that will focus on specific factors that influence the efficiency with which separately practiced tasks are integrated. Candidate factors include level of proficiency attained on the tasks before the student attempts to learn to integrate the tasks, the sequence in which the separate tasks are learned, the use of verbal mediation, and the provision of cues during practice on separate tasks that serve as "bridges" to related tasks.

The purpose of the research proposed here is (1) to formulate principles for separate task practice that incorporate guidance pertaining to the facilitation of subsequent integration of the tasks; and (2) to determine empirically the adequacy of the principles through examination of the effects of their use on student learning.

Several research projects outlined in Appendix B would involve learning analyses for particular purposes. If these efforts were completed first, investigators involved in the research proposed here would have actual examples available for guidance in deriving guides for general use.

Two kinds of validations of learning analyses would be of interest. First, recommended facets of each program developed could be compared

with existing programs to see what differences, and additional insights, resulted from learning analyses. Second, the efficiency and effectiveness of any program developed following the guides could be validated as outlined in Appendix B for demonstrating the value of an operational program, or a program for teaching selected skills.

3. TECHNIQUES FOR TEACHING ROBUST AIRCREW SKILLS

Experienced pilots can transition from one type of aircraft to another more readily than inexperienced pilots. They can more readily adapt performances to the requirements of the situation at hand. They can also adapt quickly to the peculiarities of an ATD. This adaptability is possible because, with experience, the "bandwidth" of flying skills usually increases to accommodate variations, even distortions, in cues and in response requirements. Skills characterized by the ability to make these accommodations with only temporary, minor deterioration in performance have been termed "robust skills" by some researchers.

The development of robust skills is an important goal for any career aircrewman, for usually he will transition to a number of different aircraft and have to perform under a variety of situations. The sooner his skills become robust, the less transition training he will need and the safer will be his performance.

As ATDs are used more and more in training, and especially as their use in continuation training increases, the early development of robust skills could overcome many fidelity shortcomings of the devices. Since some fidelity shortcomings will always exist in ATDs, and others can be remedied only at great cost, robust skills can provide assistance to the mediational processes. Thus, there is a need to develop techniques for teaching skill robustness, both for training efficiency and for improved safety in performance.

The purpose of the research proposed here would be to develop such techniques. A programmatic effort would eventually be involved, but the analytic portion of early projects would target the general problem in a way that subsequent efforts would reduce to validations of procedures for particular skills. It would be necessary to analyze in substantial detail dimensions of cues and responses involved in skills so that generalizable discriminations could be specifically targeted for training. In empirical tests of proposed training regimens, it would be important to recognize that negative transfer would be common, but that it is by overcoming, not necessarily avoiding, negative transfer that robustness will be acquired. Of course, disruptions in performance should be minimized, but only so long as their avoidance does not vitiate the purpose of the training.

4. MEASURES OF TRANFSER OF TRAINING

The value of ATDs rests entirely in the extent to which skills trained in them become manifest in aircraft performance. This transfer of ATD learning is a very complex process, and current measures of it do not identify facets that are crucial to valid evaluations of ATD training effectiveness and efficiency. For example, the widely used Transfer Effectiveness Ratio (TER) confuses enhanced levels of original aircraft performance upon transition from an ATD with improved rate of aircraft learning because of foundations acquired in ATDs. Furthermore, the projected maximum benefit, i.e., the asymptotic level of aircraft performance, has not even been recognized as dependent on transfer of ATD skills.

The complexity of transfer exceeds these three types of manifestation, i.e., original proficiency level in aircraft, rate of aircraft learning, and asymptotic level of performance. Nevertheless, separate reliable measures of at least the first two would be of inestimable value for training decisions and research. They would permit determination, for example, of positive versus negative transfer either as a beginning aircraft proficiency level or as a rate of aircraft learning. Had adequate measures been available, almost surely some previous research would have revealed that, in spite of favorable TERs, negative transfer on one factor, either a beginning level or a learning rate, had to be compensated for by positive transfer on the other. If it could be known reliably where the negative transfer occurs, more enlightened decisions regarding its avoidance would be possible.

This is only one example of how more precise measures could be of value. The pervasive shortcomings of available measures are evident in the wide acknowledgment that there are no measures of effectiveness of training suitable for precise cost effectiveness analyses. For this purpose, it would be necessary to estimate asymptotic performance levels in addition to the other two measures of transfer.

The purpose of the research proposed here would be to devise techniques for measuring transfer that would provide reliable estimates of (1) beginning level of aircraft proficiency upon transition from an ATD; (2) post-transition rate of learning; and (3) asymptotic performance levels. The effort would be mainly analytic, and it would include specifications for quantifying variables to assure that the derived measures had desired meanings. Empirical efforts would illustrate how the measures could be obtained and used.

5. TECHNIQUES FOR CONTROL OF THE QUALITY OF ATD TRAINING

Numerous factors can act to degrade the effectiveness of an ongoing ATD training program--e.g., turnover in instructional personnel, changes in the quality of student input, modifications of managerial practices. To prevent degradation, the effectiveness and efficiency of ATD training should be monitored routinely, and, when the quality of training is observed to decrease, changes should be made to counter the degrading factors.

As discussed in the Utilization report, the absence of a reliable mechanism for systematically monitoring ATD training effectiveness on a continuing basis was judged to be an important deficiency in the management of most of the ATD training surveyed during STRES. It was also observed that while ISD procedures explicitly require the monitoring of all ISD developed training, these procedures have yet to be extended formally to most ATD training. Furthermore, AFM 50-2 and other documents defining such procedures do not specify how monitoring of ATD training should be accomplished.

To provide guidance for the development of systems for controlling the quality of ATD training, programmatic research is needed to identify effective techniques for routinely measuring the effectiveness and efficiency of ATD training. The research proposed here would involve a substantial analytic effort that would identify (1) specific requirements for effective training quality control systems for different types of ATD training programs; (2) requirements for information concerning student and instructor performance that is needed to support such systems; (3) candidate measure sets and analysis procedures to provide the needed information; and (4) requirements for training the personnel who will use the system.

Following this analysis, a quality control system would be designed and developed for a selected ATD training program. After its implementation, the system would be operated on an experimental basis for 12 - 18 months in order to provide a basis for its evaluation. The results of such evaluation would provide a basis for the development and institution of ATD training quality control programs on a wide basis.

6. FIELD OF VIEW REQUIREMENTS

The specification of field of view (FOV) requirements for ATDs is an extremely critical design decision due to the substantial increase in direct and indirect costs associated with increasing FOV, and to the extensive impact of the decision on the number and types of tasks that can be taught in the ATD. As discussed in the Fidelity report, this decision is complicated further by differences in cue requirements related to the use of foveal and peripheral vision.

While there has been considerable research on the processing of foveally perceived information, little research has been conducted, until relatively recently, that focused on the nature and use of peripherally perceived information. This recent research has revealed that peripheral vision is involved in a large number of perceptual judgments that are extremely important to the performance of many aircrew tasks--e.g., directing foveal vision in visual search, estimation of relative velocity and of the path of motion of objects, maintenance of bodily orientation. Much is still not known, however, concerning the nature of cues that are processed via peripheral vision. Thus, while it may be relatively easy to determine cue requirements associated with the use of foveal vision, it is more difficult to identify requirements for cues that are processed peripherally.

Programmatic research is needed to identify FOV requirements associated with the instruction of different aircrew tasks. As part of this program of research, the study proposed here would determine cue requirements associated with the use of peripheral vision. The results of this research would aid designers in determining not only FOV requirements, but also area of interest display requirements.

7. EFFECTS OF ASYNCHRONIZATION OF VISUAL AND MOTION-FORCE CUES

Due to unavoidable time requirements for certain electronic and mechanical actions in ATDs, an unnatural delay can occur between a control input by a pilot and the provision of feedback to him through visual display systems and motion-force cueing devices. As discussed in the Fidelity report, two questions or problems arise. The first is, How much delay in visual or motion feedback can be tolerated before the pilot's ability to learn a task is compromised, before acceptance of a device suffers, or before illness becomes a problem? The second question is, Given the possibility of differences in the amount of delay between normally simultaneous related visual and motion-force cues, how much asynchronization in these delays can be tolerated before learning, acceptance, or illness become problems?

While there has been sufficient research on the first question to provide at least tentative guidance to designers, very little empirical data exists that can be employed to develop specific guidelines concerning cue asynchronization questions. Research is needed to develop a data base that can be employed to focus simulation technology research on the most critical aspects of the asynchronization problem, and to guide the utilization of existing ATDs with asynchronized visual and motion-force cues.

The proposed program of research addresses the problems of determining the effects of asynchronization for different training objectives. Independent variables in the research would be (1) type of training objective classified in terms of required cue and response discriminations; (2) type of asynchronization (i.e., visual lags motion versus motion lags visual); (3) degree of asynchronization; and (4) amount of delay of the visual and motion feedback. Dependent variables would assess ATD training efficiency and effectiveness of transfer to the aircraft.

8. DESIGN AND USE OF AN AUTOMATED PROCEDURES MONITORING SYSTEMS

Most automated performance measurement research has focused on the assessment of psychomotor performance. There is also a need for the capability to measure and evaluate procedural performance automatically. While the monitoring of procedures is usually a simple task for an instructor, it is also time consuming. Automation of the process would provide him with more time to devote to other instructional functions. Additionally, automated procedures monitoring would enable greater use of self-instruction, peer instruction, and instruction by non-rated personnel.

The design of an automated procedures monitoring system is complicated, however, by the existence of procedures that are not inflexible, linear sequences--e.g., procedures where steps can be performed in different orders, procedures in which certain steps may be optional depending on the situation, procedures in which any one of a set of alternative steps can be selected. Thus, automated recording, analysis, and evaluation of procedural performance can present difficult design problems. (Indeed, a major engineering research effort is needed to develop software that can handle these problems at a practical cost.)

Moreover, little is known concerning the information requirements of the instructor with respect to evaluation of procedural performance. Since the output of an automated procedures monitoring system must be useful for real-time management of the learning process, these information requirements need to be identified.

Research is needed to (1) identify candidate dimensions of procedural performance that can be measured; (2) develop means for their quantification; and (3) develop methods of evaluating and scoring the measures.

The research proposed here has a substantial analytic component that would address these issues on a general level. A subsequent experiment would also be planned in which an automated procedures monitoring system would be developed for a selected ATD and its use in training evaluated.

9. USE OF COMPUTER SPEECH RECOGNITION AND GENERATION TECHNOLOGY

Computer speech recognition and generation technology has advanced considerably in the past decade and promises to develop rapidly in the near future. Numerous possibilities exist for the application of this technology in ATD system design. For example, the computer could be employed to relieve the instructor of the burden of simulating crew or extended team members during individual training. Limitations on the simulation of these personnel by instructors now restrict this type of training. Additionally, computer speech and recognition capabilities could be used to replace the instructor and to facilitate ATD-controlled instruction for some tasks. When the instructor is present, the computer could automatically relieve him of selected coaching, cueing, and tutoring tasks. It would also be possible to use computer speech understanding systems to enable oral control of the ATD. Finally, given the use of computer speech recognition and generation systems as an interface option onboard aircraft, these systems will need to be employed in ATDs to enable practice of their use.

Research is needed now to determine if, when, and how this technology can be incorporated into ATD system design. Indeed, the use of these systems in some training devices is already being investigated by various agencies. However, given the numerous possibilities that exist for the use of computer speech capabilities, programmatic research is required to enable efficient exploration of ATD applications.

Part of the research proposed would be devoted to the analysis of computer speech recognition and generation technology to identify potential applications of the technology to specific Air Force ATD training programs. In recognition of the urgent need to improve extended team training in the ATDs, the second part of the research would (1) identify feasible applications of the technology for extended team training that are currently possible; (2) develop a computer speech recognition and generation system that could be employed in an existing ATD; and (3) evaluate the effectiveness of the use of the system for training extended team tasks.

APPENDIX D

MASTER LISTING OF POTENTIAL RESEARCH TOPICS

PRECEDING PAGE BLANK-NOT FILMED

This appendix contains the list of 111 potential research topics from which the final sets of primary and secondary topics were selected. Based on the procedures described in Chapter II, each of the topics was classified in one of three categories of research priority (denoted by the numbers "1", "2", and "3", in the column next to the topics). Topics that were judged to warrant detailed consideration were classified in the first category ("1"). The appendix, B or C, in which these topics are discussed is identified in the "comments" column next to each topic.

Topics that were deemed to be important, but not to warrant detailed consideration were classified in the second category ("2"). Generally, nothing appears in the "comments" column for these second category topics. Topics classified in the third category ("3") were judged not to warrant research. The reasons for their exclusion from research consideration are included in the "comments" column next to each of these topics.

To facilitate their review, the 111 topics were grouped into the 17 major areas shown below.

TOPICS	PRIORITY CATEGORY	COMMENTS
A. LEARNING ANALYSIS AND THE DEVELOPMENT OF ATD DESIGN AND TRAINING REQUIREMENTS		
1. Evaluation of the use of learning analysis to develop ATD design requirements and aircrew design programs.	1	Topics A.1 and A.2 are combined and discussed as one topic in Appendix C. Learning analysis is also a major concern in many of the other topics discussed in Appendices B and C.
2. Evaluation of the use of learning analysis to revise existing ATD training program.	1	
B. ORGANIZATION OF TASKS FOR PRACTICE IN ATDS		
1. Determination of the effectiveness of providing separate practice on tasks that normally occur concurrently or in succession during operational performance.	2	
2. Determination of the relationships between conditions for separate task practice and the subsequent integration of the tasks.	1	Discussed in Appendix C.
3. Evaluation of the use of "low fidelity" ATDs for providing practice on separated tasks.	2	

TOPICS	PRIORITY CATEGORY	COMMENTS
C. INDIVIDUALIZATION OF ATD INSTRUCTION		
1. Identification of effective techniques for the individualization of training in team training.	1	Combined with topic G.1 and discussed in Appendix B.
2. Assessment of transfer as a function of amount of training.	2	
3. Identification of student characteristics that should be considered in the individualization of ATD instruction.	2	
D. LOCUS OF CONTROL OF ATD INSTRUCTION		
1. Determination of effective techniques for the use of self-instruction in ATDs.	1	Discussed in Appendix B.
2. Determination of ATD design requirements for enabling and facilitating self-instruction.	2	
3. Determination of effective techniques the use of peer instruction in ATDs.	2	
4. Determination of ATD design requirements for enabling and facilitating peer instruction.	2	
5. Determination of effective techniques for the use of dynamic observation in ATDs.	2	

TOPICS	PRIORITY CATEGORY	COMMENTS
6. Determination of ATD design requirements for enabling and facilitating the use of dynamic observation.	2	
E. ATD DESIGN REQUIREMENTS AND TECHNIQUES FOR MANAGEMENT OF THE INSTRUCTIONAL PROCESS		
1. Programmatic research to determine effective techniques for the use of guidance.	1	Combined with E.3 and discussed in Appendix B.
2. Programmatic research to identify innovative ATD design features that facilitate the use of guidance.	2	
3. Programmatic research to determine effective techniques for the use of feedback.	1	Combined with E.1 and discussed in Appendix B.
4. Programmatic research to identify innovative ATD design features that facilitate the use of feedback.	2	
5. Programmatic research to determine effective techniques for the use of mediation in ATD instruction.	2	Mediation is a major consideration in much of the research discussed in Appendices B and C.
6. Programmatic research to determine effective techniques for the integrated use of feedback, guidance, and mediation.	2	

TOPICS	PRIORITY CATEGORY	COMMENTS
F. METHODS OF USE OF INSTRUCTIONAL SUPPORT FEATURES		
1. Determine effective techniques for the use of programmed mission scenarios in ATD instruction.	2	
2. Determine effective techniques for the use of programmed mission scenarios in proficiency checkrides.	2	
3. Identification of effective automated performance measures and scoring techniques for use in automated adaptive training.	2	
4. Identification of effective adaptive logics for use in automated adaptive training.	2	
5. Identification of effective techniques for using automated demonstrations, freeze, record/replay, and automated cueing and coaching in conjunction with automated adaptive training.	2	
6. Programmatic research to evaluate the use of automated adaptive training in operational ATD training programs.	2	

TOPICS	PRIORITY CATEGORY	COMMENTS
7. Identification of effective techniques for the use of total system freeze.	3	With respect to topics F.7 through F.11, it was felt that research questions should focus on instructional process variables, such as guidance and feedback, rather than be formulated in terms of specific issues concerning the use of freeze.
8. Identification of effective techniques for the use of flight systems freeze.	3	
9. Identification of effective techniques for the use of position freeze.	3	
10. Identification of effective techniques for the use of parameter freeze.	3	
11. Identification of effective techniques for the integrated use of total system, flight system, position and parameter freeze.	3	
12. Comparative evaluation of the training effectiveness of predictable versus highly interactive "iron pilots."	2	
13. Assessment of the training effectiveness of the use of "iron pilots" in relation to pilot skill level.	2	
14. Investigation of the uses of computer speech recognition and generation technology in ATDs.	1	Discussed in Appendix C.

TOPICS	PRIORITY CATEGORY	COMMENTS
15. Evaluation of the use of auto-demos as performance models.	3	Research should focus on the more general issues related to instructional process variables (see F.7).
16. Evaluation of the use of auto-demos for tactics development and dissemination.	1	The more general problem of determining how to use ATDs in tactics development and dissemination is discussed in Appendix B.
17. Evaluation of the use of auto-demos for introduction to aircraft handling characteristics.	2	
18. Evaluation of the use of auto-demos to introduce ATD features and characteristics.	2	
19. Evaluation of the use of auto-demos as a means of assessing instructor grading practices.	2	
20. Evaluation of the use of slow-time auto-demos in comparison to real-time auto-demos.	2	
21. Investigation of principles of use of record/replay.	2	
22. Determination of supplemental information that can be used to enhance the instructional value of replay.	3	It was felt that this question should be reformulated in terms of instructional process issues.
23. Determination of optimal message content and timing for automated cueing and coaching.	2	

TOPICS	PRIORITY CATEGORY	COMMENTS
24. Determination of methods for providing automated cueing and coaching.	2	
25. Exploratory research to develop and evaluate methods for automated procedures monitoring.	1	Discussed in Appendix C.
26. Development of functional requirements and ATD design technology for use of delayed remote displays of performance.	1	Discussed in Appendix B.
G. DESIGN OF ATDs AND DEVELOPMENT OF TRAINING FOR THE INSTRUCTION OF CREWS		
1. Determination of effective techniques for crew instruction.	1	Discussed in Appendix B.
2. Determination of effective ATD design features for use in crew training.	2	
3. Determination of effective techniques for "extended" team training.	1	Discussed in Appendix B.
H. DESIGN OF ATDs AND DEVELOPMENT OF TRAINING FOR HIGH AROUSAL TASKS		
1. Identification of effective techniques for use in teaching tasks that are normally performed under conditions of high arousal.	2	

TOPICS	CATEGORY	COMMENTS
2. Determination of effective ATD design features for use in teaching high arousal tasks.	2	
I. ATD INSTRUCTOR FACTORS		
1. Determination of requirements for ATD instructor training programs.	1	Discussed in Appendix B.
2. Comparison of alternative (a) techniques for structuring the management of the ATD instructional process, and (*b*) *methods* of training the ATD instructor to use those techniques.	2	
3. Determination of effective techniques of using ATDs in ATD and flight instructor training.	2	
4. Development and evaluation of innovative ATD design features that facilitate the use of ATDs in instructor training.	2	
5. Evaluation of the use of a centrally produced, audio-visual program for initial or in-service training of ATD instructors.	2	
6. Identification of characteristics of candidate ATD instructors that predict the proficiency of their subsequent instructional performance.	2	

TOPICS	PRIORITY CATEGORY	COMMENTS
7. Evaluation of the use of an ATD instructor selection system.	2	
8. Determination of effective techniques for evaluating ATD instructor proficiency.	1	Discussed in Appendix B.
J. PERFORMANCE MEASUREMENT IN ATDs		
1. Determination of effective techniques for assessing trainee performance.	1	Discussed in Appendix B.
2. Determination of effective techniques for assessing team performance.	1	Discussed in Appendix B.
3. Development of measures of transfer of training.	1	Discussed in Appendix C.
4. Determination of effective measures and reliable procedures for evaluating the effectiveness of ATD training organizations.	2	
5. Assessment of the use of training quality control systems in ATD training programs.	1	Discussed in Appendix C.
6. Determination of ATD design requirements for the use of ATDs in proficiency evaluations.	1	Discussed in Appendix B.

TOPICS	PRIORITY CATEGORY	COMMENTS
K. CUE ANALYSIS AND THE DEVELOPMENT OF ATD DESIGN REQUIREMENTS		
1. Identification of cues that aircrewmen actually use in operational performance.	2	
2. Development of a lexicon for describing cue system (e.g., visual) requirements that would replace the use of "fidelity" as a descriptor.	2	
3. Determination of effective techniques for teaching robust aircrew skills.	1	Discussed in Appendix C.
4. Identification of interactions between cues that impact on learning.	2	
5. Identification of the minimum cues necessary for training.	2	
L. MOTION-FORCE CUE REQUIREMENTS		
1. Determination of motion and force cue requirements.	1	Discussed in Appendix B.
2. Evaluation of alternative mechanisms for providing motion-force cues.	1	Discussed in Appendix B.
3. Use of g-seats for disturbance motion cues.	2	Included in L.2

TOPICS	PRIORITY CATEGORY	COMMENTS
4. Determination of the effectiveness of use of ALCOGs-type g-seats for presenting vertical axis (heave) cues.	2	Included in L.2
M. VISUAL CUE REQUIREMENTS		
1. Investigation of specific parameters of visual displays (e.g., resolution, luminance, color, delays, level of detail).	2	Research on the specific details of scene content is discussed in Appendix B. This research would cover many of the topics listed in Section M.
2. The size, density and number of embedded levels of texture necessary to afford perception of a solid surface and perspective transformation information for ground speed, altitude, and distance should be investigated.	2	
3. Field of view requirements (FOV) for training of different flight tasks should be determined.	1	Discussed in Appendix C.
4. Relative importance and effects of perceptual learning in a flight simulator on subsequent flying of the aircraft should be investigated.	2	

TOPICS	PRIORITY CATEGORY	COMMENTS
5. The characteristics of the borders of dynamically positionable area of interest (AOI) displays should be investigated	2	
6. The amount of detail required to represent objects sufficiently in simulated scenes should be investigated.	2	
7. Investigation of the contribution to training effectiveness of the use of infinity optics.	2	
8. Determination of visual cue requirements.	1	Discussed in Appendix B.
9. Evaluation of the use of visual displays to present disturbance cues.	2	
N. IOS DESIGN		
1. Determination of the most effective position for ATD instructors for the conduct of ACM training.	1	Discussed in Appendix B.
2. Evaluation of alternative display strategies for presenting the external visual scene to the instructor.	1	Discussed in Appendix B.
3. Development of measures of instructor workload criteria for IOS design.	2	

TOPICS	PRIORITY CATEGORY	COMMENTS
O. DESIGN OF ATDs AND DEVELOPMENT OF OPERATIONAL TRAINING PROGRAMS		
1. Identification of effective training techniques for teaching selected operational tasks.	1	Topics O.1 and O.2 are combined and discussed in Appendix B.
2. Identification of ATD design features for use in teaching selected operational tasks.	1	
3. Evaluation of the systematic application of training technology to the development of an operational aircrew (ATD) training program.	1	Topics O.3 and O.4 are combined and discussed in Appendix B.
4. Evaluation of the systematic application of training technology to the revision of an operational aircrew (ATD) training program.	1	
P. DESIGN OF ATDs AND DEVELOPMENT OF TRAINING FOR COMPLEX COGNITIVE SKILLS		
1. Identification of effective techniques for use in training complex cognitive skills.	1	Discussed in Appendix B.
2. Identification of ATD design features that facilitate teaching complex cognitive skills.	2	

TOPICS	PRIORITY CATEGORY	COMMENTS
Q. MISCELLANEA		
1. Evaluation of the utility of the worth of ownership model presented in the Worth of Ownership volume.	3	A management issue.
2. Development of procedures for the collection of cost data associated with the use of ATDs.	3	A management issue.
3. Identification of effective ways to teach selection of pertinent cues and responses.	2	
4. Determination of how and when it is possible to use an ATD for a given aircraft for training for another aircraft.	2	
5. Evaluation of the effectiveness, in terms of transfer of training, of making the ATD training task more difficult than the task to be performed in the aircraft.	2	
6. Determine the need for standardization of flight training programs by equipment type within the Air Force.	3	A management issue.
7. Research to determine effective methods for facilitating the update of ATDs.	2	

TOPICS	PRIORITY CATEGORY	COMMENTS
8. Development and validation of a model for predicting ATD training effectiveness.	1	Discussed in Appendix B.
9. Development of criteria that can be used by instructors to determine if a partially malfunctioning ATD can be used for training.	2	
10. Investigate and define criteria for ATD acceptance with respect to stick response and instrument movement, particularly with high activity training sessions.	3	Q.10 and Q.11 do not focus on the real issue of training value.
11. Determine the significance of stick force fidelity by measuring the stick force on similar equipment such as the UPT-IFS. Verify if the stick force in each ATD reacts in the same or different fashions and investigate the complaints or comments by the users on stick force.	3	
12. Evaluation of the effectiveness of the SAAC/TAC-ACES program	3	Research is not necessary.
13. Identification of pilot strategies in aircraft control and determination of effective instructional techniques for use in their instruction.	2	
14. Identification of effective techniques for teaching spatial orientation.	2	

TOPICS	PRIORITY CATEGORY	COMMENTS
15. Determination of effective ATD design features for teaching spatial orientation.	3	Research is not needed.
16. Determination of the effects of asynchronization of visual and motion-force cues.	1	Discussed in Appendix C.
17. Investigation of techniques for preparing students for ATD training sessions.	2	
18. Assessment of the effects of session length on the efficiency and effectiveness of ATD instruction.	3	Length per se should not be the focus of training research.
19. Determination of effective techniques for integrating ATD instruction with training in other media.	2	

DATE
FILMED

8